

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)  
14423.488 Million cell updates/sec

14423.488 Million cell updates/sec

Sequence: 1 tgcacaacatgtccaaaaa...gtagttattaaaaaaa 920

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Minimum Match	0%
Maximum Match	100%

Listing first 45 summaries

Database :

1:	qb_bmb1.*
2:	qb_bmb.*
3:	qb_hdg.*
4:	qb_in.*
5:	qb_om.*
6:	qb_ov.*
7:	qb_pat.*
8:	qb_ph.*
9:	qb_pl.*
10:	qb_pr.*
11:	qb_ro.*
12:	qb_sbs.*
13:	qb_sy.*
14:	qb_un.*
15:	qb_v1.*
16:	em_ba.*
17:	em_fun.*
18:	em_in.*
19:	em_in.*
20:	em_om.*
21:	em_or.*
22:	em_ov.*
23:	em_pat.*
24:	em_pl.*
25:	em_ro.*
26:	em_sbs.*
27:	em_sy.*
28:	em_un.*
29:	em_v1.*
30:	em_hdg_bum.*
31:	em_hdg_inv.*
32:	em_hdg_other.*
33:	em_hdg_mus.*
34:	em_hdg_pln.*
35:	em_hdg_pod.*
36:	em_hdg_nam.*
37:	em_hdg_vrt.*
38:	em_sy.*
39:	em_hngo_bum.*
40:	em_hngo_mus.*
41:	em_hngo_other.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	920	100.0	920	6	A16257	i16257 Synthetic R
2	920	100.0	920	6	A16258	i16257 Synthetic R
3	920	100.0	920	6	AR080406	AR080406 Sequence
4	920	100.0	920	6	AR092530	AR092530 Sequence
5	920	100.0	920	6	AR122885	AR122885 Sequence
6	920	100.0	920	6	AR123540	AR123540 Sequence
7	920	100.0	920	6	AR148357	AR148357 Sequence
8	911.8	99.1	923	14	HRSRNAG	X17085 Human resp
9	909.2	98.8	922	14	HRSVGL16	Z33429 Human resp
10	901.6	98.0	917	14	HRSHGLY	M17212 Human resp
11	894	97.2	894	6	AR080424	AR080424 Sequence
12	894	97.2	894	6	AR092548	AR092548 Sequence
13	853.4	92.8	8510	14	RSNLE	M11486 Human resp
14	853.4	92.8	15222	6	AR093219	AR093219 Sequence
15	853.4	92.8	15222	14	HRU50363	U50363 Human resp
16	853.4	92.8	15222	14	HRU50363	U50363 Human resp
17	853.4	92.8	15222	14	HRU63644	U63644 Human resp
18	853.4	92.8	15222	14	RSNLS	M74568 Human resp
19	853.4	92.8	15223	6	AR089137	AR089137 Sequence
20	853.4	92.8	15223	14	AF035006	AF035006 Human res
21	848.2	92.2	914	14	PARSENVG	X03149 Respiratory
22	848.2	92.2	918	6	AX047075	AX047075 Sequence
23	837.8	91.1	897	6	AX339011	AX339011 Sequence
24	837.2	91.0	15190	14	HRU38662	U39662 Human resp
25	835.6	90.8	15191	14	RSU38661	U39661 Respiratory
26	820.4	89.2	918	14	AF065406	AF065406 Human res
27	814	88.5	918	14	AF065405	AF065405 Human res
28	811.6	88.2	922	14	HRSCVG1	Z33410 Human resp
29	811.6	88.2	922	14	HRSCVGA	Z33454 Human resp
30	810	88.0	922	14	HRSCVGA	Z33412 Human resp
31	810	88.0	922	14	HRSCVGC	Z33456 Human resp
32	810	88.0	922	14	HRSVGL1	Z33424 Human resp
33	810	88.0	922	14	HRSVGL2	Z33425 Human resp
34	808.4	87.9	922	14	HRSVGL11	Z33494 Human resp
35	808.4	87.9	922	14	HRSVGL12	Z33427 Human resp
36	807.2	87.7	922	14	HRSVGL14	Z33425 Human resp
37	806.8	87.7	922	14	HRSVGL2	Z33415 Human resp
38	806.8	87.7	922	14	HRSVGL9	Z33422 Human resp
39	805.2	87.5	922	14	AF065256	AF065256 Human res
40	805.2	87.5	922	14	HRSVGL10	Z33423 Human resp
41	805.2	87.5	922	14	HRSVGL13	Z33426 Human resp
42	805.2	87.5	922	14	HRSVGL3	Z33416 Human resp
43	803.6	87.3	922	14	AF065255	AF065255 Human res
44	803.6	87.3	922	14	AF065258	AF065258 Human res
45	803.6	87.3	922	14	HRSVGL17	Z33430 Human resp



QY	601	CACACCCAAAGCTACAAAAAAACCAAGCTCTAAGACACACAAAAAAGATCTCTAAACCTCA	660
Db	320	CACCAACCAAGCTACAAAAAAACCAAGCTCTAAGACACACAAAAAAGATCTCTAAACCTCA	261
QY	661	AACCACTAAACCCAAAGGAAGTACCCACCACACAGCCACAGAAAGACCAACCATCAACAC	720
Db	260	AACCACTAAACCCAAAGGAAGTACCCACCACACAGCCACAGAAAGACCAACCATCAACAC	201
QY	721	CACCAAAACAAACATACACACTACAGCTCTACACCAACACACACAGAAATCCAAAACCT	780
Db	200	CACCAAAACAAACATACACACTACAGCTCTACACCAACACACACAGAAATCCAAAACCT	141
QY	781	CACAAGTCAAAATGGAACCTTCCACTCAACCTCCTCCGAAGCAATCTAAGCCCTTCTCA	840
Db	140	CACAAGTCAAAATGGAACCTTCCACTCAACCTCCTCCGAAGCAATCTAAGCCCTTCTCA	81
QY	841	AGTCCGCCAACATCCGAGCAACCCATCACAACCCATCTCCACCCACACACAACAGCGCA	900
Db	80	AGTCCGCCAACATCCGAGCAACCCATCACAACCCATCTCCACCCACACACAACAGCGCA	21
QY	901	GTAGTTATTTAAAAAAGAAAAA 920	
Db	20	GTAGTTATTTAAAAAAGAAAAA 1	

[illegible]

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 920)	Klein, M.H., and Ewaszshyn, M.E.	Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus

JOURNAL	Patent:	US 5968776-A	7	19-Oct-1999;
FEATURES	Location/Qualifiers			
Source	1..920			
	/organism="unknown"			
BASE COUNT	380 a	290 c	95 g	155 t
ORIGIN				
Query Match	100.0%;	Score 920:	DB 6;	Length 920;
Best Local Similarity	100.0%;	Pred. No. 3,6e-190;		
Matches 920; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Db	301	AATCAGCTTCTCCAAATCTGTCTGAAATTAACATCCAAACACACCACCATATCTAGCTTCAAC	360
OY	361	AATCAGCAGAGTCAAATCTCAAACTGCAACCCCAACAGTCGAAGACTTAAACACAAACAC	420
Db	361	AAACCCAGGAGTCGAATCTCAAACTTGCACCCCAACAGTCGAAGACTTAAACACAAACAC	420
OY	421	AAACCCAAACACAAACCCGAGAGCCCACTACAAACCAAGCCCAAAACCAACAC	480
Db	421	AAACCCAAACACAAACCCGAGAGCCCACTACAAACCAAGCCCAAAACCAACCAAC	480
OY	481	ACCCAAATATATGATTTTCACTTGGAGAGTGTTTAACTTTGTACCTCGAGCATATGAGACAA	540
Db	481	ACCCAAATATATGATTTTCACTTGGAGAGTGTTTAACTTTGTACCTCGAGCATATGAGACAA	540
OY	541	CAATCCAAACCTCTGGGCTATCTGCAAAAGATACCAAAACAAACCAACGAGAAAGAAAAC	600
Db	541	CAATCCAAACCTCTGGGCTATCTGCAAAAGATACCAAAACAAACCAACGAGAAAGAAAAC	600
OY	601	CACCAACCAAGCTTCAAAAAAACCAACCTTCAAGACACCAAAAAAGATCTCAAACTTCA	660
Db	601	CACCAACCAAGCTTCAAAAAAACCAACCTTCAAGACACCAAAAAAGATCTCAAACTTCA	660
OY	661	AAACCAATTAACCAAGAGAGTACCAACACCAAGCCACAGAGAGCCCAACATTAATAC	720
Db	661	AAACCAATTAACCAAGAGAGTACCAACACCAAGCCACAGAGAGCCCAACATTAATAC	720
OY	721	CACCAAAACCAAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACT	780
Db	721	CACCAAAACCAAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACT	780
OY	781	CACCAATCAAAATGGAACCTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACT	840
Db	781	CACCAATCAAAATGGAACCTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACTTCCAACT	840
OY	841	AGCTTCACAATATCCGAGACCACTTCAACACCACTTCAATTCACCAACCAACAAACAGGCA	900
Db	841	AGCTTCACAATATCCGAGACCACTTCAACACCACTTCAATTCACCAACCAACCAACAGGCA	900
OY	901	GTAAGTATTTAAAAAAGAAAAA 920	
Db	901	GTAAGTATTTAAAAAAGAAAAA 920	

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	GI:10019284	DNA	linear	PAT 08-SEP-2000
AR092530	AR092530	Sequence 7 from patent US 5998169.	AR092530	AR092530.1	GI:10019284			

REFERENCE AUTHORS TITLE	unclassified. 1 (bases 1 to 920) Klein, M. H., Du, R.-P. and Ewasysbyn, M. E. Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial
-------------------------------	---

```

VIRUS
Patent: US 5998169-A 7 07-DEC-1999;
JOURNAL
Location/Qualifiers
FEATURES
1. .920
source
/organism="unknown"
BASE COUNT
380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Prid. No. 3.ee-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TGCAAACATGTTCCAAAACAAAGACCAACGGACCGCTTAGACACTAGAAAAGACCTGGGA 60  
|||||  
1 TGCAAACATGTTCCAAAACAAAGACCAACGGACCGCTTAGACACTAGAAAAGACCTGGGA 60  
Db 1 TGCAAACATGTTCCAAAACAAAGACCAACGGACCGCTTAGACACTAGAAAAGACCTGGGA 60

QY 61 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120  
 Db 61 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120  
 QY 121 AGCAAAATCATATTATTCATTTCTGGCAATGATATATCTCACTTCTTAATTTACAGC 180  
 Db 121 AGCAAAATCATATTATTCATTTCTGGCAATGATATATCTCACTTCTTAATTTACAGC 180  
 QY 181 CATCATTTATATAGCTTCGGCAAAACCAAGTCACTTAACAGTCAATCTATACAGA 240  
 Db 181 CATCATTTATATAGCTTCGGCAAAACCAAGTCACTTAACAGTCAATCTATACAGA 240  
 QY 241 TGCACACAGCCAGATCAAGAACACACCCCACTACTCTAGAGTCTTCTGAGTCTTG 300  
 Db 241 TGCACACAGCCAGATCAAGAACACACCCCACTACTCTAGAGTCTTCTGAGTCTTG 300  
 QY 301 AATAGCTTCTCAATCTGTCTGAATTTATCATCAACACCAACCACTACTACTTCAAC 360  
 Db 301 AATAGCTTCTCAATCTGTCTGAATTTATCATCAACACCAACCACTACTACTTCAAC 360  
 QY 361 AACACAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAACTCAACACAC 420  
 Db 361 AACACAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAACTCAACACAC 420  
 QY 421 AACCCAAACACACCCGCAAGCCCACTACAAAACACCCCAACACACCCCAACAA 480  
 Db 421 AACCCAAACACACCCGCAAGCCCACTACAAAACACCCCAACACACCCCAACAA 480  
 QY 481 ACCCAATATGATTTTACCTTCAAGTGTGTTAACTTTGACCTGCGAGATATGACGAA 540  
 Db 481 ACCCAATATGATTTTACCTTCAAGTGTGTTAACTTTGACCTGCGAGATATGACGAA 540  
 QY 541 CAATCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAAC 600  
 Db 541 CAATCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAAC 600  
 QY 601 CACACCAAGCTCAAAAAAACCACTTCAAGACAAACCAAAAAAATCTCAAACTCA 660  
 Db 601 CACACCAAGCTCAAAAAAACCACTTCAAGACAAACCAAAAAAATCTCAAACTCA 660  
 QY 661 AACCACTAAACCAAGAGTACCCACACCAACCCCAAGAGAGGCAACCATCAACAC 720  
 Db 661 AACCACTAAACCAAGAGTACCCACACCAACCCCAAGAGAGGCAACCATCAACAC 720  
 QY 721 CACCAAAACCAACATCACTACATCTGCTACCAACCAACCAAGAAATCCAAACT 780  
 Db 721 CACCAAAACCAACATCACTACATCTGCTACCAACCAACCAAGAAATCCAAACT 780  
 QY 781 CACAAGTCAAAATGGAACCTTCCACTACCTCTCGGAAGGCAATTAAGCCCTTCTCA 840  
 Db 781 CACAAGTCAAAATGGAACCTTCCACTACCTCTCGGAAGGCAATTAAGCCCTTCTCA 840  
 QY 841 AGTCTCACAACATCGAGCACCCATCAACCTCTCTCAACCAACCAACACAGCCCA 900  
 Db 841 AGTCTCACAACATCGAGCACCCATCAACCTCTCTCAACCAACCAACACAGCCCA 900  
 QY 901 GTAGTTATTTAAAAAAA 920  
 Db 901 GTAGTTATTTAAAAAAA 920

RESULT 5  
 ARI22885  
 LOCUS ARI22885 920 bp DNA linear PAT 16-MAY-2001  
 DEFINITION Sequence 7 from patent US 6168786.  
 ACCESSION ARI22885  
 VERSION ARI22885.1 GI:14107851  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 920)  
 AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyn, M.E.

TITLE Chimeric immunogens  
 JOURNAL Patent: US 6168786-A 7 02-JAN-2001;  
 FEATURES location/Qualifiers  
 source 1..920  
 BASE COUNT 380 a 290 c 95 g 155 t  
 ORIGIN  
 Query Match 100.0%; Score 920; DB 6; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 3,6e-190;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAACATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAAACCTGGGA 60  
 Db 1 TGCAAACATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAAACCTGGGA 60  
 QY 61 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120  
 Db 61 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120  
 QY 121 AGCAAAATCATATTATTCATTTCTGGCAATGATATATCTCACTTCTTAATTTACAGC 180  
 Db 121 AGCAAAATCATATTATTCATTTCTGGCAATGATATATCTCACTTCTTAATTTACAGC 180  
 QY 181 CATCATTTATATAGCTTCGGCAAAACCAAGTCACTTAACAGTCAATCTATACAGA 240  
 Db 181 CATCATTTATATAGCTTCGGCAAAACCAAGTCACTTAACAGTCAATCTATACAGA 240  
 QY 241 TGCACACAGCCAGATCAAGAACACACCCCACTACTCTAGAGTCTTCTGAGTCTTG 300  
 Db 241 TGCACACAGCCAGATCAAGAACACACCCCACTACTCTAGAGTCTTCTGAGTCTTG 300  
 QY 301 AATAGCTTCTCAATCTGTCTGAATTTATCATCAACACCAACCACTACTACTTCAAC 360  
 Db 301 AATAGCTTCTCAATCTGTCTGAATTTATCATCAACACCAACCACTACTACTTCAAC 360  
 QY 361 AACACAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAACTCAACACAC 420  
 Db 361 AACACAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAACTCAACACAC 420  
 QY 421 AACCCAAACACACCCGCAAGCCCACTACAAAACACCCCAACACACCCCAACAA 480  
 Db 421 AACCCAAACACACCCGCAAGCCCACTACAAAACACCCCAACACACCCCAACAA 480  
 QY 481 ACCCAATATGATTTTACCTTCAAGTGTGTTAACTTTGACCTGCGAGATATGACGAA 540  
 Db 481 ACCCAATATGATTTTACCTTCAAGTGTGTTAACTTTGACCTGCGAGATATGACGAA 540  
 QY 541 CAATCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAAC 600  
 Db 541 CAATCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAAC 600  
 QY 601 CACACCAAGCTCAAAAAAACCACTTCAAGACAAACCAAAAAAATCTCAAACTCA 660  
 Db 601 CACACCAAGCTCAAAAAAACCACTTCAAGACAAACCAAAAAAATCTCAAACTCA 660  
 QY 661 AACCACTAAACCAAGAGTACCCACACCAACCCCAAGAGAGGCAACCATCAACAC 720  
 Db 661 AACCACTAAACCAAGAGTACCCACACCAACCCCAAGAGAGGCAACCATCAACAC 720  
 QY 721 CACCAAAACCAACATCACTACATCTGCTACCAACCAACCAAGAAATCCAAACT 780  
 Db 721 CACCAAAACCAACATCACTACATCTGCTACCAACCAACCAAGAAATCCAAACT 780  
 QY 781 CACAAGTCAAAATGGAACCTTCCACTACCTCTCGGAAGGCAATTAAGCCCTTCTCA 840  
 Db 781 CACAAGTCAAAATGGAACCTTCCACTACCTCTCGGAAGGCAATTAAGCCCTTCTCA 840  
 QY 841 AGTCTCACAACATCGAGCACCCATCAACCTCTCTCAACCAACCAACACAGCCCA 900  
 Db 841 AGTCTCACAACATCGAGCACCCATCAACCTCTCTCAACCAACCAACACAGCCCA 900  
 QY 901 GTAGTTATTTAAAAAAA 920



```
Db 901 GTAGTTATTAAAAAAA 920
|||||
RESULT 6
LOCUS AR123540 920 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 7 from patent US 6171783.
ACCESSION AR123540
VERSION AR123540.1 GI:14108901
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Infection detection method using chimeric protein
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.6e-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAACATGTGCAAAAACAGGACCAACGCCGCTAAGACTAGAAAAAGCCTGGGA 60
|||
Db 1 TGCAAACATGTGCAAAAACAGGACCAACGCCGCTAAGACTAGAAAAAGCCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
|||
Db 61 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCACAATCACATTATTCATTCTGGCAATGATATCTCAACTCATTATTAATTCAGC 180
|||
Db 121 AGCACAATCACATTATTCATTCTGGCAATGATATCTCAACTCATTATTAATTCAGC 180
QY 121 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
|||
Db 181 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
QY 181 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
|||
Db 181 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
QY 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
|||
Db 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
QY 301 AATCAGCTTCTGCAATCTGTCTGAAATTACATCACAAAACCAACCATACTAGCTTCAAC 360
|||
Db 301 AATCAGCTTCTGCAATCTGTCTGAAATTACATCACAAAACCAACCATACTAGCTTCAAC 360
QY 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAAGCTCAAGATTAATAAACACAGCAAC 420
|||
Db 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAAGCTCAAGATTAATAAACACAGCAAC 420
QY 421 AACCCAAACACAGCCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACCAAA 480
|||
Db 421 AACCCAAACACAGCCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACCAAA 480
QY 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTGTACCTGACAGCATATGAGCAA 540
|||
Db 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTGTACCTGACAGCATATGAGCAA 540
QY 541 CAATCCAACTGCTGGCTATCTGCAAAAAGATACAAACAAAAAACCCAGAAAGAAAAC 600
|||
Db 541 CAATCCAACTGCTGGCTATCTGCAAAAAGATACAAACAAAAAACCCAGAAAGAAAAC 600
QY 601 CACCAACAAGCCTACAAAAAACCAAGCTTCAAGACACCAAAAAAGATCTCAAACTCA 660
|||
Db 601 CACCAACAAGCCTACAAAAAACCAAGCTTCAAGACACCAAAAAAGATCTCAAACTCA 660
QY 661 AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGCCACAGAGCAACATCAACAC 720
|||
```

```
Db 661 AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGCCACAGGAAGGCAATCAACAC 720
|||||
QY 721 CACCAAAACAAATCAATCAATCTGCTGCTACCAACACACACAGGAAATCCAAACT 780
|||
Db 721 CACCAAAACAAATCAATCAATCTGCTGCTACCAACACACACAGGAAATCCAAACT 780
QY 781 CACAAGTCAAAATGGAACCTTCCACTGACCTGCTGGAAGGCAATCTAGCCCTTCTCA 840
|||
Db 781 CACAAGTCAAAATGGAACCTTCCACTGACCTGCTGGAAGGCAATCTAGCCCTTCTCA 840
QY 841 AGTCTCCACAACATCCGAGCACCATCAACACCTCATCTCACTCCACCAACACAGGCCA 900
|||
Db 841 AGTCTCCACAACATCCGAGCACCATCAACACCTCATCTCACTCCACCAACAGCAGGCCA 900
QY 901 GTAGTTATTAAAAAAA 920
|||
Db 901 GTAGTTATTAAAAAAA 920
|||||
RESULT 7
LOCUS AR148357 920 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 7 from patent US 6225091.
ACCESSION AR148357
VERSION AR148357.1 GI:15112447
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers
protection against parainfluenza virus and respiratory syncytial
virus
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.6e-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAACATGTGCAAAAACAGGACCAACGCCGCTAAGACTAGAAAAAGCCTGGGA 60
|||
Db 1 TGCAAACATGTGCAAAAACAGGACCAACGCCGCTAAGACTAGAAAAAGCCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
|||
Db 61 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCACAATCACATTATTCATTCTGGCAATGATATCTCAACTCATTATTAATTCAGC 180
|||
Db 121 AGCACAATCACATTATTCATTCTGGCAATGATATCTCAACTCATTATTAATTCAGC 180
QY 181 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
|||
Db 181 CATCATATTCATAGGCTGGGCAACGACCAAGCTCACTAATCACTGCAATCATACAGA 240
QY 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
|||
Db 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
QY 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
|||
Db 241 TGCACAAGCCGATCAAGAACACACACCCCAATACCTCACTCGAGATCCTCAGCTTGG 300
QY 301 AATCAGCTTCTGCAATCTGTCTGAAATTACATCACAAAACCAACCATACTAGCTTCAAC 360
|||
Db 301 AATCAGCTTCTGCAATCTGTCTGAAATTACATCACAAAACCAACCATACTAGCTTCAAC 360
QY 361 AACCAACAAGGAGTCAAGTCAAACTGCAACCCCAAGCTCAAGATTAATAAACACAGCAAC 420
|||
Db 361 AACCAACAAGGAGTCAAGTCAAACTGCAACCCCAAGCTCAAGATTAATAAACACAGCAAC 420
|||
```



RESULT 9  
HRSVGL16 922 bp RNA linear VRL 05-JUN-1997  
LOCUS Human respiratory syncytial virus (MON-7-91) subgroup A, G gene for  
DEFINITION glycoprotein.  
ACCESSION 233429.1 GI:485888  
VERSION 233429.1  
KEYWORDS G gene; glycoprotein.  
SOURCE Human respiratory syncytial virus.  
ORGANISM Human respiratory syncytial virus.  
Viruses: ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Dopazo, J.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-1994) Dopazo, J., Centro Nacional de Biotecnologia  
- CSIC, Bioinformatica, Universidad Autonoma, Cantoblanco, Madrid,  
SPAIN, 28049  
REFERENCE 2 (bases 1 to 922)  
AUTHORS Garcia, O., Martin, M., Dopazo, J., Arbiza, J., Fabrasile, S., Russi, J.,  
Hortal, M., Perez-Brena, P., Martinez, I., Garcia-Barreno, B. and  
Melero, J. A.  
TITLE Evolutionary pattern of human respiratory syncytial virus (subgroup  
A): cocirculating lineages and correlation of genetic and antigenic  
changes in the G glycoprotein  
JOURNAL J. Virol. 68 (9), 5448-5459 (1994)  
MEDLINE 9435057  
PUBMED 8057427  
FEATURES  
source location/Qualifiers  
1..922  
/organism="Human respiratory syncytial virus"  
/strain="subgroup A"  
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"  
/db\_xref="taxon:11250"  
/cell\_line="HEp-2"  
16..912  
/gene="G"  
16..912  
/partial  
/gene="G"  
/codon\_start=1  
/product="glycoprotein"  
/protein\_id="CAA83872.1"  
/db\_xref="GI:485889"  
/translation="ISKNDQPTAKTLEKTDUNLILFISGLYKLNKSLAQITLS  
ILAMITSLIITAIIFIASANKHVTLLAIODASQIKNTPTLTDDPOLGISFS  
NLSEITISQTTILASTTPGKSNLOPTTYKTKTTTOTOPEKPTIKOKNPKPNEN  
NDHFEVFNPCISCSNNPCMAICKRIIPNKKPKKTKPKPKPKTKKDKLPQ  
TKPKREVPITKPEPTINTTKITTTITLTNTNTPKLTLSOMETFEHSTISSEGNLSP  
SOVSTISEHSPSPSPNTTRQ"  
BASE COUNT 378 a 291 c 97 g 156 t  
ORIGIN  
Query Match 98.8%; Score 909.2; DB 14; Length 922;  
Best Local Similarity 99.7%; Pred. No. 8e-188;  
Matches 911; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

241 TGCACAGCCAGATCAAGACACACACCCACATACCTACTCAGATCTCAGCTTGG 300  
|||||  
Db 249 TGCACACAGCCAGATCAAGACACACACCCACATACCTACTCAGATCTCAGCTTGG 308  
|||||  
Qy 301 AATCAGCTTCCAAATCTGTCTGAAATATACATCACAAACCCACCACATCTAGCTTCAAC 360  
|||||  
Db 309 AATCAGCTTCCAAATCTGTCTGAAATATACATCACAAACCCACCACATCTAGCTTCAAC 368  
|||||  
Qy 361 AACACGAGGTAAGTCAAACTGCAACCCCAACAGTCAAGACTTAAACATCAAC 420  
|||||  
Db 369 AACACGAGGTAAGTCAAACTGCAACCCCAACAGTCAAGACTTAAACATCAAC 428  
|||||  
Qy 421 AACCCAAACACACCCAGACGCCACTACAAAACACGCCCAAAACACCAACCA 480  
|||||  
Db 429 AACCCAAACACACCCAGACGCCACTACAAAACACGCCCAAAACACCAACCA 488  
|||||  
Qy 481 ACCCAATATATATTTTCACTTGGAAGTGTTAACCTTTGACCTGACATATGACGCA 540  
|||||  
Db 489 ACCCAATATATATTTTCACTTGGAAGTGTTAACCTTTGACCTGACATATGACGCA 548  
|||||  
Qy 541 CAATCCACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCCGAAACAAAC 600  
|||||  
Db 549 CAATCCACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCCGAAACAAAC 608  
|||||  
Qy 601 CACCACCAAGCCTACAAAAAAACCAACCTTCAAGACACCAAAAAAGATCTCAAACTCA 660  
|||||  
Db 609 CACCACCAAGCCTACAAAAAAACCAACCTTCAAGACACCAAAAAAGATCTCAAACTCA 668  
|||||  
Qy 661 AACCACTTAAACCAAGGAAGTACCCACACCAAGCCGACAGAAAGCCACATCAACAC 720  
|||||  
Db 669 AACCACTTAAACCAAGGAAGTACCCACACCAAGCCGACAGAAAGCCACATCAACAC 728  
|||||  
Qy 721 CACCAAAACCAACATACAACTACACTGTCACCAACACACACGAGAAATCCAAACT 780  
|||||  
Db 729 CACCAAAACCAACATACAACTACACTGTCACCAACACACACGAGAAATCCAAACT 788  
|||||  
Qy 781 CACAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAATCTTCTCA 840  
|||||  
Db 789 CACAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTTAAAGCCCTTCTCA 848  
|||||  
Qy 841 AGTCTCCACACATCTCGAGACCCATACACACCTCTATCTCCACCCACACACGCCCA 900  
|||||  
Db 849 AGTCTCCACACATCTCGAGACCCATACACACCTCTATCTCCACCCACACACGCCCA 908  
|||||  
Qy 901 GTAGTTATTAATAA 914  
|||||  
Db 909 GTAGTTATTAATAA 922  
|||||  
RESULT 10  
RSHGLYG 917 bp ss-RNA linear VRL 03-AUG-1993  
LOCUS Human respiratory syncytial virus (subgroup A) attachment protein  
DEFINITION (G) mRNA, complete cds.  
ACCESSION M17212  
VERSION M17212.1 GI:333940  
KEYWORDS attachment glycoprotein; surface glycoprotein.  
SOURCE Human respiratory syncytial virus (subgroup A, strain Long), cDNA to  
viral RNA, clones pL63, pL69, and pA3.  
ORGANISM Human respiratory syncytial virus  
Viruses: ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 917)  
AUTHORS Johnson, P. R., Springs, M. K., Olmsted, R. A. and Collins, P. L.  
TITLE The G glycoprotein of human respiratory syncytial viruses of  
subgroups A and B: extensive sequence divergence between  
antigenically related proteins  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)  
MEDLINE 87289657  
PUBMED 2441388  
COMMENT The exact 5' end of Long G mRNA was not determined.  
FEATURES location/Qualifiers



Db 481 AATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGAGCATATGAGCAACAATGCA 540  
 QY 548 ACCTGTGGGCTATCTGGAAAAGATTACCAACAAAACCGGAAAGAAACACCGACC 607  
 Db 541 ACCTGTGGGCTATCTGGAAAAGATTACCAACAAAACCGGAAAGAAACACCGACC 600  
 QY 608 AAGCTTACAAAAAACCAACCTTCAAGACAAACAAAAAGATCTCAAACTCAAACT 667  
 Db 601 AAGCTTACAAAAAACCAACCTTCAAGACAAACAAAAAGATCTCAAACTCAAACT 660  
 QY 668 AAACCAAGGAAGTACCCACCAACGACCAAGGCAAGGCAACCATCAGACCAACCA 727  
 Db 661 AAACCAAGGAAGTACCCACCAACGACCAAGGCAAGGCAACCATCAGACCAACCA 720  
 QY 728 ACAACATCACAACCTACCTGCTGACCAACCAACCAAGCAATCCAAACTCAACAAGT 787  
 Db 721 ACAACATCACAACCTACCTGCTGACCAACCAACCAAGCAATCCAAACTCAACAAGT 780  
 QY 788 CAATGGAAACCTTCACACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847  
 Db 781 CAATGGAAACCTTCACACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840  
 QY 848 ACAACATCCGAGACCCATCCACACCCCTCATCTCCACCAACCAACCAACGCGCAG 901  
 Db 841 ACAACATCCGAGACCCATCCACACCCCTCATCTCCACCAACCAACCAACGCGCAG 894

RESULT 12  
 AR092548  
 LOCUS AR092548 894 bp DNA linear PAT 08-SEP-2000  
 DEFINITION Sequence 28 from patent US 5998169.  
 ACCESSION AR092548  
 VERSION AR092548.1 GI:10019302  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 894)  
 AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyh, M.E.  
 TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
 JOURNAL Patent: US 5998169-A 28 07-DEC-1999;  
 FEATURES Location/Qualifiers  
 source 1..894  
 BASE COUNT 364 a 288 c 93 g 149 t

Query Match 97.2%; Score 894; DB 6; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-184;  
 Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATGTCCAAAACAAAGACCAACGACCGCTAAGACACTAGAAAAGACGCGACACTCTC 67  
 Db 1 ATGTCCAAAACAAAGACCAACGACCGCTAAGACACTAGAAAAGACGCGACACTCTC 60  
 QY 68 AATCATTTTATTTATTCATATCATCGGGCTTATTAATTAACTTAACTTGTAGACAA 127  
 Db 61 AATCATTTTATTTATTCATATCATCGGGCTTATTAATTAACTTAACTTGTAGACAA 120  
 QY 128 ATTCATTTTATTCATTTGTGCAATGATATCTCAACTACTATATATACGCAATCAT 187  
 Db 121 ATTCATTTTATTCATTTGTGCAATGATATCTCAACTACTATATATACGCAATCAT 180  
 QY 188 TTTCATAGCTCGGCAACCAAGTACAGTAACTGCAATCATACAGATGACAA 247  
 Db 181 TTTCATAGCTCGGCAACCAAGTACAGTAACTGCAATCATACAGATGACAA 240  
 QY 248 AACGAGATCAAGACCAACCAACATATCATCTACATGAGATCTTCAAGTATGACG 307  
 Db 241 AACGAGATCAAGACCAACCAACATATCTTACATGAGATCTTCAAGTATGACG 300

QY 308 TTTCGAATCTGTCTGAATTTACATCAACAAACCAACCATACTAGCTTCACAGACCA 367  
 Db 301 TTTCGAATCTGTCTGAATTTACATCAACAAACCAACCATACTAGCTTCACAGACCA 360  
 QY 368 GGAGTCAAGTCAAAACCGCAACCCACAGAGTCAAGCTTAAAAACACACACAAACCGCA 427  
 Db 361 GGAGTCAAGTCAAAACCGCAACCCACAGAGTCAAGCTTAAAAACACACACAAACCGCA 420  
 QY 428 ACACACCCAGCAAGCCCTACTACAAAAACAAGCCAAACCAACCATCAAAACCAAT 487  
 Db 421 ACACACCCAGCAAGCCCTACTACAAAAACAAGCCAAACCAACCATCAAAACCAAT 480  
 QY 488 AATGATTTTACCTTGAAGTGTAACTTTGTACCTGACGATATGCAAGCTTCAATGCA 547  
 Db 481 AATGATTTTACCTTGAAGTGTAACTTTGTACCTGACGATATGCAAGCTTCAATGCA 540  
 QY 548 ACCTGTGGGCTATCTGGAAAAGATTACCAACAAAACCGGAAAGAAACACCGACC 607  
 Db 541 ACCTGTGGGCTATCTGGAAAAGATTACCAACAAAACCGGAAAGAAACACCGACC 600  
 QY 608 AAGCTTACAAAAAACCAACCTTCAAGACAAACAAAAAGATCTCAAACTCAAACT 667  
 Db 601 AAGCTTACAAAAAACCAACCTTCAAGACAAACAAAAAGATCTCAAACTCAAACT 660  
 QY 668 AAACCAAGGAAGTACCCACCAACGACCAAGGCAAGGCAACCATCAGACCAACCA 727  
 Db 661 AAACCAAGGAAGTACCCACCAACGACCAAGGCAAGGCAACCATCAGACCAACCA 720  
 QY 728 ACAACATCACAACCTACCTGCTGACCAACCAACCAAGCAATCCAAACTCAACAAGT 787  
 Db 721 ACAACATCACAACCTACCTGCTGACCAACCAACCAAGCAATCCAAACTCAACAAGT 780  
 QY 788 CAATGGAAACCTTCACACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847  
 Db 781 CAATGGAAACCTTCACACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840  
 QY 848 ACAACATCCGAGACCCATCCACACCCCTCATCTCCACCAACCAACCAACGCGCAG 901  
 Db 841 ACAACATCCGAGACCCATCCACACCCCTCATCTCCACCAACCAACCAACGCGCAG 894

RESULT 13  
 RSHICE 8510 bp ss-RNA linear VRL 29-NOV-2000  
 LOCUS  
 DEFINITION Human respiratory syncytial virus nonstructural protein (1c), nonstructural protein (1b), major nucleocapsid (N), phosphoprotein (P), protein (M), 1A (1A), G (G), protein (F) and envelope-associated protein (22K) gene, complete cds.  
 ACCESSION M11486 K01459 K02719 K03348 K03349 M11247 M11244 M11487 M11505  
 VERSION M11486.1 GI:333925  
 KEYWORDS envelope-associated protein; fusion glycoprotein; major nucleocapsid protein; major surface glycoprotein; matrix protein; nonstructural protein; phosphoprotein.  
 SOURCE Human respiratory syncytial virus.  
 ORGANISM Human respiratory syncytial virus.  
 Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 REFERENCE 1 (bases 1085 to 2220; 5254 to 5544)  
 AUTHORS Elango, N. and Venkatesan, S.  
 TITLE Amino acid sequence of human respiratory syncytial virus nucleocapsid protein  
 JOURNAL Nucleic Acids Res. 11 (17), 5941-5951 (1983)  
 MEDLINE 83299261  
 PUBMED 6310521  
 REFERENCE 2 (bases 3211 to 4157)  
 AUTHORS Satake, M. and Venkatesan, S.  
 TITLE Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein  
 JOURNAL J. Virol. 50 (1), 92-99 (1984)  
 MEDLINE 84138836  
 PUBMED 6699948  
 REFERENCE 3 (bases 2288 to 3191)

AUTHORS	Strake, M., Eliango, N. and Venkatesan, S.
TITLE	Sequence analysis of the respiratory syncytial virus phosphoprotein gene
JOURNAL	J. Virol. 52 (3), 991-994 (1984)
MEDLINE	85033973
PUBMED	6548527
REFERENCE	4 (bases 5602 to 7500)
AUTHORS	Collins, P.L., Huang, Y.T. and Wertz, G.W.
TITLE	Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)
MEDLINE	85088471
PUBMED	6096849
REFERENCE	5 (bases 7551 to 8510)
AUTHORS	Collins, P.L. and Wertz, G.W.
TITLE	The envelope-associated 22k protein of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polytranscript
JOURNAL	J. Virol. 54 (1), 65-71 (1985)
MEDLINE	85135082
PUBMED	3838351
REFERENCE	6 (bases 8 to 1050; 7554 to 8506)
AUTHORS	Eliango, N., Strake, M. and Venkatesan, S.
TITLE	mRNA sequence of three respiratory syncytial virus genes encoding two nonstructural proteins and a 22k structural protein
JOURNAL	J. Virol. 55 (1), 101-110 (1985)
MEDLINE	85237684
PUBMED	4009789
REFERENCE	7 (bases 5602 to 7423)
AUTHORS	Eliango, N., Strake, M., Colligan, J.E., Norrby, E., Camargo, E. and Venkatesan, S.
TITLE	Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit
JOURNAL	Nucleic Acids Res. 13 (5), 1559-1574 (1985)
MEDLINE	85215565
PUBMED	2987829
REFERENCE	8 (bases 4630 to 5543)
AUTHORS	Satake, M., Colligan, J.E., Eliango, N., Norrby, E. and Venkatesan, S.
TITLE	Respiratory syncytial virus envelope glycoprotein (G) has a novel structure
JOURNAL	Nucleic Acids Res. 13 (21), 7795-7812 (1985)
MEDLINE	86067198
PUBMED	4069997
REFERENCE	9 (bases 4627 to 5544)
AUTHORS	Wertz, G.W., Collins, P.L., Huang, Y., Gruber, C., Levine, S. and Ball, L.A.
TITLE	Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 82 (12), 4075-4079 (1985)
MEDLINE	85216636
PUBMED	3858865
REFERENCE	10 (bases 4173 to 4571)
AUTHORS	Collins, P.L. and Wertz, G.W.
TITLE	The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript
JOURNAL	Virology 141 (2), 283-291 (1985)
MEDLINE	86098645
PUBMED	3879976
REFERENCE	11 (bases 1 to 528; 552 to 1050)
AUTHORS	Collins, P.L. and Wertz, G.W.
TITLE	Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus
JOURNAL	Virology 143 (2), 442-451 (1985)
MEDLINE	86045805
PUBMED	2998021
REFERENCE	12 (bases 1081 to 2277)
AUTHORS	Collins, P.L., Anderson, K., Langer, S.J. and Wertz, G.W.
TITLE	Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus
JOURNAL	Virology 146 (1), 69-77 (1985)
MEDLINE	85301874
PUBMED	3839952

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)	Collins, P.L., Dickens, L.E., Buckler-White, A., Olmsted, R.A., Spriggs, M.K., Camargo, E. and Coelingh, K.V.	Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)	86259643	3460060	Reprint for [2] and clean copy sequence for [12], [5] kindly provided by P. Collins, 21-FEB-1986.
						Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown.
FEATURES	SOURCE	Location/Qualifiers				
gene	CD5	1..8510	/organism="Human respiratory syncytial virus"			
gene	CD5	1..8510	/strain="A2"			
gene	CD5	1..8510	/db_xref="taxon:11250"			
gene	CD5	1..8510	/lab_host="Hep-2 cell"			
gene	CD5	1..8510	/gene="1C"			
gene	CD5	1..8510	/gene="1C"			
gene	CD5	1..8510	/gene="1C"			
gene	CD5	1..8510	/codon_start=1			
gene	CD5	1..8510	/product="nonstructural protein"			
gene	CD5	1..8510	/protein_id="AAB59850.1"			
gene	CD5	1..8510	/db_xref="GI:333926"			
gene	CD5	1..8510	/translation="MGSSLSMIKVRIONLFNDNEVALIKTCYDKLIHLTNLAARAGVHTIKLNGIVFVHTTSSDICTPNNIVYKSNFTTMTPLONGGIWEMMELTHCSQNRGLIDNCEIKRFSKSLSDSTMTNMQNSSELLGFDLNP"			
variation	variation	171				
variation	variation	171	/gene="1C"			
variation	variation	171	/note="c in [11]; t in [6]"			
variation	variation	171	/replace="c"			
variation	variation	171	552..1050			
variation	variation	171	/gene="1B"			
variation	variation	171	552..1050			
variation	variation	171	/gene="1B"			
variation	variation	171	584..958			
variation	variation	171	/gene="1B"			
variation	variation	171	/codon_start=1			
variation	variation	171	/product="nonstructural protein"			
variation	variation	171	/protein_id="AAB59851.1"			
variation	variation	171	/db_xref="GI:333927"			
variation	variation	171	/translation="MDTTHNDNTQQRIMITDMPISLETTITSLTRDITTKFYILINHECIVRKLDKQAFVTFPLVEMKILKRVGSTRKKRYEYNTKGTGTFMPPLFINHDSLECIDKPTKTHPIYKYDLNP"			
variation	variation	171	747			
variation	variation	171	/gene="1B"			
variation	variation	171	/note="a in [11]; g in [6]"			
variation	variation	171	/replace="g"			
variation	variation	171	769			
variation	variation	171	/gene="1B"			
variation	variation	171	/note="g in [11]; a in [6]"			
variation	variation	171	/replace="a"			
variation	variation	171	1081..2277			
variation	variation	171	/gene="N"			
variation	variation	171	1081..2277			
variation	variation	171	/gene="N"			
variation	variation	171	1096..2271			
variation	variation	171	/gene="N"			
variation	variation	171	/codon_start=1			
variation	variation	171	/product="major nucleocapsid"			

```

/protein_id="AA559852.1"
/db_xref="GI:333928"
/translation="MALSKVKLNDLTINKDQLSSKYYTITQSTGDSITDPNYDQKAT
NKLGGKLTIEDANRKTGTGIQMLYAMSRUGREDITIKLRDAGYVKNAGVDYTHRC
DINGEMKEMFEVLTIASLTTEITQIMLIEESRKSYKKILRMGEVAPYHRDSDGMITI
LCIALVHTYIKAAADBSRGSLVAIVIRRNANLKNMKRYKGLKLLPQIDANSFYVFEKHP
FIDVHEVHGISTQRCGSGVEGIFAGLPMNAYGACQVLTBPGYI LAKSYVNIIGHSN
VOAEKQVQVYEVYEAOKLGSGEAFYHILNNRKASLSLTQPHSSSVLVGNAAGLITN
GEYKTRPNRDVLTIDAAATVAEOLKENGVINYSVLDTAEELIATIKHOLPKNDIVEL"
/gene="p"
/2289..3196
/gene="p"
/2289..3196
/gene="p"
/302..3027
/gene="p"
/godon_start=1
/product="Phosphoprotein"
/protein_id="Aab59853.1"
/db_xref="GI:333929"
/translation="MEKEAFEGEDDANNRATKPLESTIKGKTSKDPKKKSDITSYV
SIDELVESKPTISNSNIINPTREDDTQCNPNQYRKPLYSFKDPPTSPNPSPKLT
KETLEFDNNKEESSYSYSEINQTNIDNTAKLRIDELSLTIGLMTLLVAASGPTL
SARGIDANLALBKEETIGRLTIDELMTNDRLEAARLRNESEKMAKDTSDVSLNMP
ISEKLNLTGEGNSDNDLSLEDE"

```

Query Match	92.8%;	Score 853.4;	DB 14;	Length 8510;
Best Local Similarity	95.5%;	Pred. No. 9.6e-176;		
Matches 878;	Conservative	0;	Mismatches 41;	Indels 0;
				Gaps 0;

OY	1	TGAAACGTGTCCAAAAACAAGAGCCAAAGCACCCTGTAAGACACTTGGAAAAAGACCTGGGA	60
Db	4634	TGCAAAACATGTGCCAAAAACAAGAGCCAAAGCACCCTGTAAGACACTTGGAAAAAGCCTGGGA	4693
OY	61	CACCTCAGTCATCTTTATTTATTCATATCATGCGGGCTTATATAGTTAAATCTTAAATCTGT	120
Db	4694	CACCTCAGTCATCTTTATTTATTCATATCATGCGGGCTTATATAGTTAAATCTTAAATCTGT	4755
OY	121	AGCAACAATCATATTATTCATTTCTGGCAATGATTAATCTCACTGACTTATTAATTAACG	180
Db	4754	AGCAACAATCATATTATTCATTTCTGGCAATGATTAATCTCACTGACTTATTAATTAACG	4813
OY	181	CATCATATTCATTAACCTCTGGGCAACCAACAAGTCACATCTACACATGTCAGTATACAAGA	240
Db	4814	CATCATATTCATTAACCTCTGGGCAACCAACAAGTCACACCAACACGTCAGTATATACAAGA	4873
OY	241	TGCAACAAGCAGATCAAGAACACAAACCCCAATATCTCACTGAGATCCTGACCTTGG	300
Db	4874	TGCAACAAGCAGATCAAGAACACAAACCCCAATATCTCACTGAGATCCTGACCTTGG	4933
OY	301	AATGAGCTGTCCAAATCTGTCTGAATTTACATACAAACCCACCATCTACTAGCTTCAAC	360
Db	4934	AATGAGCTGTCCAAATCTGTCTGAATTTACATACAAATCACCAACCATCTACTAGCTTCAAC	4993
OY	361	AACACGAGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGACTTAAAAACACAAAC	420
Db	4994	AACACGAGAGTCAAGTCAAAACCTGCAATTCACACAGTCAAGAGCCCAAAAAACACAAAC	5053
OY	421	AACCAACAACCAACCCAGCAAGGCCACTGCAAAACAAGCCCAAAACAACCAACCAACAA	480
Db	5054	AACCAACAACCAACCCAGCAAGGCCCAACCAAAACAAGCCCAAAACAACCAACCAAGCA	5113
OY	481	ACCCAAATATGATTTTACCTTGCAGAGTGTTTAACTTTGTACCTGCAGCATATGCAGCA	540
Db	5114	ACCCAAATATGATTTTACCTTGTGAAAGTGTTCAACTTTGTACCTGCAGCATATGCAGCA	5173
OY	541	CAATCCAACTGCTGGGGCTATCTGCAAAAGATTAACAAACAAAACCAAGAAAGAAAC	600
Db	5174	CAATCCAACTGCTGGGGCTATCTGCAAAAGATTAACAAACAAAACCAAGAAAGAAAC	5233
OY	601	CACACACAACTTACAAAAAACCAACTTTCAGACAAACCAAAAAAGATCTCAAACTCA	660
Db	5234	CACACACAACTTACAAAAAACCAACTCTTCAGAACCAACCAAAAAAGATCTCCAAACTCA	5293

QY	661	AACACTAAACCAAAGAGTACCACACACACCAAGCCACAGAAAGGCCAACATCAACAC	720
Db	5294	AACCACTAAATTCAAAGGAGTACCCACACCAAGGCCACAGAAAGGCCAACATCAACAC	5353
QY	721	CACCAAAACAAACATCACTACATGCTGCACCAACACACACACAGAAATCCAAAAC	780
Db	5354	CACCAAAACAAACATCATATCACTACTACTCACTCCACACACACAGAAATCCAGAACT	5413
QY	781	CACAAGTCAATGGAAACCTTCACCTCAACCTCTCCGGAAGCAATCTAAGCCCTTCTCA	840
Db	5414	CACAAGTCAATGGAAACCTTCACCTCAACCTCTCCGGAAGCAATCTCAAGCCCTTCTCA	5473
QY	841	AGCTCCACACATCCGACACCCATCAACACCCATATCTCCACCCACACACACAGCGCA	900
Db	5474	AGCTCTACACACATCCGAGTACCCATCAACACCTTATCTCCACCCACACACAGCGCA	5533
QY	901	GTTAGTTATTAAAAAAA 919	
Db	5534	GTTAGTTATTAAAAACATA 5552	
RESULT	14		
LOCUS	AR093219	15222 bp	DNA
DEFINITION	Sequence 23 from patent US 5998602.		linear
ACCESSION	AR093219		
VERSION	AR093219.1	GI:10019970	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 15222)		
AUTHORS	Torrence,P.F., Silverman,R.Hugh., Clirno,N.Mario., Li,G. and Xiao,W.		
TITLE	Rnase L activators and antisense oligonucleotides effective to treat RSV infections		
JOURNAL	Patent: US 5998602-A 23 07-DEC-1999;		
FEATURES	Location/Qualifiers		
source	1..15222		
BASE COUNT	5924 a 2704 c 2356 g 4238 t		
ORIGIN			
Query Match	92.8%;	Score 853.4;	DB 6;
Best Local Similarity	95.5%;	Pred. No. 9.2e-176;	Length 15222;
Matches 878;	Conservative 0;	Mismatches 41;	Indels 0;
			Gaps 0;
QY	1	TGCAAAACATGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA	60
Db	4681	TGCAAAACATGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA	4740
QY	61	CACCTGCAATCATTTATTATTCATATCATATCGGGCTTATATAGTTAAATCTTAATCTGT	120
Db	4741	CACCTGCAATCATTTATTATTCATATCATATCGGGCTTATATAGTTAAATCTTAATCTGT	4800
QY	121	AGCAAAATGCATATATTCATATTCGCGCAATGATATATCTCAACTTCACCTTATATTACAGC	180
Db	4801	AGCAAAATGCATATATTCATATTCGCGCAATGATATATCTCAACTTCACCTTATATTGAGC	4860
QY	181	CATCATATTTCATAGCGTCGGCAACCAACAAGTACACATAACACTGCAATCATATACAGA	240
Db	4861	CATCATATTTCATAGCGTCGGCAACCAACAAGTACACATAACACTGCAATCATATACAGA	4920
QY	241	TGCAACAGGACGATCAAGAACACACACCCACACATACCTCAGATCCTCAGCTTGG	300
Db	4921	TGCAACAGGACGATCAAGAACACACCCACACATACCTCAGATCCTCAGCTTGG	4980
QY	301	AATCAGCTTTCACATCTGCTGAAATTTAATCTCAACACCAACGATCTAGCTTCAAC	360
Db	4981	AATCAGTCCCTCTATATCGCTGTGAATTTACATCACAATACCAACCTACTAGCTTCAAC	5040
QY	361	AACACAGGAGTCAAGTCAAACTGCAACCCACAACAGTCAAGCTTAAAAACACACAAAC	420

LOCUS	AR093219	15222 bp	DNA	linear	PAT 08-SEP-2000
DEFINITION	Sequence 23 from patent US 5998602.				

ACCESSION	AR093219
VERSION	AR093219.1
GI	10019970

**KEYWORDS** unknown.

ORGANISM Unknown.  
Unclassified.

REFERENCE	1 (bases 1 to 15222)
AUTHORS	Torrence, P.F., Silverman, R. Hugh., Cirino, N. Mario., Li, G. and

TITLE  
Xiao, W.  
Rfase L activators and antisense oligonucleotides effective to

Journal of  
treat RSV infections  
Patent: US 5998602-A 23 07-DEC-1999;

FEATURES	Location/Qualifiers
source	1. .15222

```

/organism="unknown"
BASE COUNT      5924 a      2704 c      2356 g      4238 t

```

# ORIGIN

Query Match	92.8%;	Score 853.4;	DB 6;	Length 15222;
Best Local Similarity	95.5%;	Pred. No. 9	2e-176.	

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCATACATGTCAAAAACAAGGACCACGACCCGCTAAGACACTAGAAAGAGACCTGGGA 60

Db 4681 TGCAACATGTCCAAAAACAAGGACCACGACCCGCTAAGACATTAGAAAGGACCTGGGA 4740

61 CACTCTCAATCATTTATTTATTCATATCATCGGCGTTATATTAAGTTAAATCTTAAATCTGT 120

Db 4741 CACCTCAATCATTATTAATTCATATCATCGTGTATATAAGTTAAATCTTAAATCTGT 4800

QY 121 AGACAAATCACATTATCCATTCTGGCAATGATTAATCTCAACTTCACTTATTAATTACAGC 180

Db 4801 AGCACAAATCACATTATCCATTCTGGCAATGATTAATCTCAACTTCACCTTATAATTGCAGC 4860

181 CATCATATTTCATAGCCCTGGCAACACAAAGTTCACACTACACACTGCAATCATACAGA 240

Db 4861 CATCATATTCATAGCCCTCGGCAACACCAAGTCACACCAACACACTGCAATCATTCAGAGA 4920

241 TGCACACAGCATCAGACACACACCCCAACATACCTCACTCAGCATCCTCAGCTTGG 300

Db 4921 TGCACACAGCCAGATCTACAAGACACAAACCCCAACATACCTTCACCCAGAGATCCTCAGCTTGG 4980

301 AATCAGCTTCTCCAATCTGTCTGAATTACATCACAACCAACCATATCTAGCTTCAAC 360

Db 4981 AATCAGTCCCTTAATCCGTGTAATTACATCACAATCACCACCACTACTAGCTTCAC 5040

361 AACACCAGAGCTCAAGTCAAACTGCAACCCACACACAGTCAAGAGCTAAAAACACACCAAC 420

```

Db      5041 AACACGAGAGTCATCAACCCCTGCATCCAAACAGTCAGACCAACCAACAAACAAAC 5100
QY      421 AACCAAACCAACCCAGCAGACCCCACTCAAAAACACGCCCAAAACCAACCAACCA 480
Db      5101 AACTCAAAACCAACCCAGCAGACCCCAACCAACCAACCAACCAACCAACCAACCA 5160
QY      481 ACCCAATTAATGATTTTCACTGCACTGTTTAACTTTTACCCCTGCAGCATATGCAAGCA 540
Db      5161 ACCCAATTAATGATTTTCACTGCACTGTTTAACTTTTACCCCTGCAGCATATGCAAGCA 5220
QY      541 CAATCCAAACCTGCTGGGCTATCTGCAAAAGAAATATCAAAACCAAAACCAAGAAAGAAAC 600
Db      5221 CAATCCAAACCTGCTGGGCTATCTGCAAAAGAAATATCAAAACCAAAACCAAGAAAGAAAC 5280
QY      601 CACCAACCAACCTGCTGGGCTATCTGCAAAAGAAATATCAAAACCAAAACCAAGAAAGAAAC 660
Db      5281 CACTACCAACCCCAACCAAAACCAACCCCTCAAGACACCAACCAAAAGATCCCAACCTCA 5340
QY      661 AACCACTAAACCAAGAAAGTACCCACCAAGCCCAAGCAAGAAAGCAACCATCAACAC 720
Db      5341 AACCACTAAATCAAGAAAGTACCCACCAAGCCCAAGCAAGAAAGCAACCATCAACAC 5400
QY      721 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCAAAACT 780
Db      5401 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCAAAACT 5460
QY      781 CACAAGCAATGAAAGAACTTCACTCAACCTCTCGAAGGCAATCTAAGCCCTCTCA 840
Db      5461 CACAAGCAATGAAAGAACTTCACTCAACCTCTCGAAGGCAATCTAAGCCCTCTCA 5520
QY      841 AGTCTCCAAACATCCGAGCAGCCATCAACCCCTCATCTCCACCAACCAAGCAAGCAAGCA 900
Db      5521 AGTCTCCAAACATCCGAGCAGCCATCAACCCCTCATCTCCACCAACCAAGCAAGCAAGCA 5580
QY      901 GTAGTTATTAAAAAAA 919
Db      5581 GTAGTTACTTAAAAACATA 5599

RESULT 15
HRU50362      15222 bp      RNA      linear      VRL-20-NOV-1997
LOCUS      Human respiratory syncytial virus, mutant cp-RSV, complete genome.
DEFINITION      U50362
ACCESSION      U50362.1 GI:2627296
VERSION      1
KEYWORDS      Human respiratory syncytial virus.
SOURCE      Human respiratory syncytial virus.
ORGANISM      Parainflaviridae; Pneumovirinae; Pneumovirus.
REFERENCE      1 (bases 1 to 15222)
AUTHORS      Connors, M., Crowe, J.E. Jr., Firestone, C.Y., Murphy, B.R. and
              Collins, P.L.
TITLE      A cold-passaged, attenuated strain of human respiratory syncytial
              virus contains mutations in the F and L genes
JOURNAL      Virology 208 (2), 478-484 (1995)
MEDLINE      95266253
PUBMED      7747420
REFERENCE      2 (bases 1 to 15222)
AUTHORS      Crowe, J.E. Jr., Firestone, C.Y., Whitehead, S.S., Collins, P.L. and
              Murphy, B.R.
TITLE      Acquisition of the ts phenotype by a chemically mutagenized
              cold-passaged human respiratory syncytial virus vaccine candidate
              results from the acquisition of a single mutation in the polymerase
              (L) gene
JOURNAL      Virus Genes 13 (3), 269-273 (1996)
MEDLINE      97187925
PUBMED      9035372
REFERENCE      3 (bases 1 to 15222)
AUTHORS      Whitehead, S.S.
TITLE      Direct Submission
JOURNAL      Submitted (29-FEB-1996) RVS, LID, NIAID, Bldg 7, Rm 118, 7 Center
              Drive, MSC 0720, Bethesda, MD 20892, USA

```

```

FEATURES
source
location/Qualifiers
1..15222
/organism="Human respiratory syncytial virus"
/strain="A2"
/db_xref="taxon:11250"
/note="mutant cp-RSV"
45..576
/gene="NS1 (1C)"
45..576
/gene="NS1 (1C)"
/product="nonstructural protein NS1"
99..518
/gene="NS1 (1C)"
/codon_start=1
/product="nonstructural protein NS1"
/db_xref="GI:2627297"
/translation="MGSNSLSMIAKYRLQNLFDNDVALLKICQYDNLHLTNALAKA
VIHRTKNGIVFVHTTSDICPNNINIVKSNFTPTPVLQNGYIWEAMELTRCSOPN
GLDDNCEIKFSKRLSDSTWTNMYNQLSELLGFPLNP"
596..1098
/gene="NS2 (1B)"
596..1098
/gene="NS2 (1B)"
/product="nonstructural protein NS2"
628..1002
/gene="NS2 (1B)"
/codon_start=1
/product="nonstructural protein NS2"
/db_xref="GI:2627298"
/translation="MDTTHNDNTPORLMTDMRPLSEITITSLRDIITHKRTIYIN
HECTVRLDERQATFPLVNEKMLHKVGSTYKRYTEYNITGYFPMPDIFINHDGF
LECGIRPKHTPIIKYDLNP"
1125..2327
/gene="N"
1125..2327
/gene="N"
/product="major nucleocapsid"
1140..2315
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/db_xref="GI:2627299"
/translation="MALSKYKNDLTUKDQLSSKTYTORSTGSDTPPNVYORHI
NKLGLMLITEDANHKFTGLIGMLYASRGREDITKILDAGIHYKANKVDYTTBRO
DINKEMKFEVLTLASLTLEIOLNIEIESRSYKMKLEGEVAPERYRHSPOGMIT
LCIALVITKILAADRSGLTAVIRANNVINKERYKGLPLDIOLANSFEVEKHPH
FIDVFVHGKIAOSSTRGSRVEGIFAGLFNNAYGAGVMDRMWGLANSININMGAS
VOAMEOVVEVEYEAOKLGGAGFYHLNPKASLSLTQEPHSSVVLGNAGLGIM
GEYRGTPRNDPLYNAKAKAYAEQLEKNENVINSVLDLAELEIAIKHQLNPKNDVEL"
2329..3242
/gene="P"
2329..3242
/gene="P"
/product="phosphoprotein"
2346..3071
/gene="P"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB86658.1"
/db_xref="GI:2627300"
/translation="MEKRAPPEFHGEDANNRATKFLSICKGFTSPDKPKKDSIYSN
SIDLEVTRESPTNSITINFTNETDTAGKPKPYOKRPLVSEKEDPTSPDNPSTKY
KETIETPDNNEESSYSYEINDOTNDITAKDRIDEKSEIIGHIHTLVASAGPT
SARDGIRDAWGLREMEIEKIRTELALMTNDLEAMRLNEESEKMAKDTSDDEVSLNP
TSEKINLLEGNDSNDLSEDF"
3252..4209
/gene="M"
3252..4209
/gene="M"
/product="Matrix protein M"

```



```
CDS
3261..4031
/gene="M"
/codon_start=1
/product="matrix protein M"
/protein_id="AAB86660.1"
/db_xref="GI:2627301"
/translation="METYVNRKLHEGSTYAAVOYNLEKDDPDASTITVWPFQSSMP
ADLLKELANVNLVROISTPKGPSLRVNSRAVLQMPKPTICANVLSDESKL
AVDTWPEKACSLTCLSKMLTIVKDIINLPHDIILALCEFNIVTSKVIIL
PYLRISVRANKDNLNTITTEFRNATINAKIIPYSGLLVITVTNKAFAKVIKP
QSQFVDCAYLEKESIVYVTINMHTATRAIKPMD"
3970..4197
/gene="M"
/note="putative alternative translation of M gene"
/codon_start=1
/product="unknown"
/protein_id="AAB86661.1"
/db_xref="GI:2627302"
/translation="MLPQIGSTQLHDLQSNPKINLPLHQCIVNSYKLTYLHFTIT
INTLWFNOSNKTYSLSQIIPSHCLSDLVK"
4219..4628
/gene="SH (1A)"
4219..4628
/gene="SH (1A)"
/product="SH (1A)"
/product="small hydrophobic protein"
/gene="SH (1A)"
/codon_start=1
/product="small hydrophobic protein"
/protein_id="AAB86662.1"
/db_xref="GI:2627303"
/translation="WENTSTIEFSKFWPYFTLLHMITTIISLIISIMAILNKL
CEYNFHNKTEFLPRARVNT"
4673..5595
/gene="G"
4673..5595
/gene="G"
/product="G"
/product="attachment G glycoprotein"
4688..5584
/gene="G"
/codon_start=1
/product="attachment G glycoprotein"
/protein_id="AAB86663.1"
/db_xref="GI:2627304"
/translation="MSKRRDQRTAKLEERTWDTLNLHLFISCLYKLNLSVAQITLS
ILAMITSLIIAIIETASANKVPTPTAIIODATSOINNTPTLYLTONPOLGISPS
NPSETTSQITIIASTPGVKSTIQTSTVTKNTTQTOPSPKTRORONKPKSPN
NDPHEVENFVPCISGNNPCMAICKRIPNKRGGKTTTKPKKPTIKTKKDDPRPO
TTKSEVETTKPTEPTINTTKNIIITLLSNTGNPELTSQMETFHSTISSEGNPS
SOVSTSEYPSPPSPNTPRO"
5648..7550
/gene="F"
5648..7550
/gene="F"
/product="F"
/product="fusion F glycoprotein precursor"
5661..7385
/gene="F"
/codon_start=1
/product="fusion F glycoprotein precursor"
/protein_id="AAB86664.1"
/db_xref="GI:2627305"
/translation="MELLILKANAITTILTAVFPCFASGQNTTEFYOSTCAVSKGY
LSALRTGWTSVITIELSNTEKKNCTDAKVKYLKQELDKRNAVTELOLMOSTPA
EGEVNKLKSLALSTNKAIVLSNGSVLSKRVRLFGFLVGSAIASGVAASKVLH
TVLEPQOQNNRLEITREFSVNAGVTTPSTYMLTUSELSIINDPITINDOKLSMN
NVOIVROQSYSIMSIIKEEVLAVVOLPLVGVLDPCMKLHSLPCTTNTKGSNICK
TRDRGMYCDNAGSVSPFOAETCKVOSNRVCDTINSLTISEVNLGVNDIFNPKYD
CKIMTSKTDVSSSVITSIGAIVSCYKTKCSNKRKGITKTFSGCDIVSNKGVDTV
SGNGLIYVWKQSGKSIYKGEPIINFTDPLVPSDEFPASTISQVNEKINSLAFIRK
SDELHIVNAGKSTINIMITTIITVITLISLIAVGLLYCARSTPVTLSKQDLSG
INNIAFSN"
7597..8557
gene
```

```
mrna
7597..8557
/gene="M2 (22K)"
/product="matrix protein M2"
7606..8190
/gene="M2 (22K)"
/codon_start=1
/product="matrix protein M2"
/protein_id="AAB86665.1"
/db_xref="GI:2627306"
/translation="MSRRNPKFEIRGHCLNKRCHFSNYFFMPHALLVRONFMJN
RLUKSDKSIDFLSETSGAALDRTEYEAALGVGLSEYGCINNTIKOSACVMSKL
LTELNDSDIKLRDNEELNSPKIRVYNTVIVSYIESNRNNKOTIHLRLPADVLKKT
IKNTLIDHKSITINPKESTVSDTMDHAKNNDTT"
8159..8431
/gene="M2 (22K)"
/note="putative alternative translation of M2 gene"
/codon_start=1
/product="unknown"
/protein_id="AAB86666.1"
/db_xref="GI:2627307"
/translation="MTMPKIMLPKYPGCSITSILITSRCRVTYNKNLTLYFNQNP"

Query Match 92.8% Score 853.4: DB 14: Length 15222;
Best Local Similarity 95.5%: Pred. No. 9,2e-176;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAAACATGCCAAAAAGACGACGACCGCTAGACACTAGAAAGACCTGGGA 60
Db 4681 TGCAAACATGCCAAAAAGACGACGACCGCTAGACACTAGAAAGACCTGGGA 4740
QY 61 CACTGCAATCATTATATATATATATATATATATATATATATATATATATAT 120
Db 4741 CACTGCAATCATTATATATATATATATATATATATATATATATATATATAT 4800
QY 121 AGCACAATACATATATATATATATATATATATATATATATATATATATATAT 180
Db 4801 AGCACAATACATATATATATATATATATATATATATATATATATATATATAT 4860
QY 181 CATCATATATATATATATATATATATATATATATATATATATATATATATATAT 240
Db 4861 CATCATATATATATATATATATATATATATATATATATATATATATATATATAT 4920
QY 241 TGCACAACCCAGATCAACAACAACCAACCCAGATCAACAACAACCCAGATCAACA 300
Db 4921 TGCACAACCCAGATCAACAACAACCAACCCAGATCAACAACAACCCAGATCAACA 4980
QY 301 AATCAGCTTCTCAATCTGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
Db 4981 AATCAGCTTCTCAATCTGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5040
QY 361 AACACGAGAGTCAGTCAAACTGCAACCCGACGACGACGACGACGACGACGACGAC 420
Db 5041 AACACGAGAGTCAGTCAAACTGCAACCCGACGACGACGACGACGACGACGACGAC 5100
QY 421 AACCCCAACACAAACCCGACGACGACGACGACGACGACGACGACGACGACGACGAC 480
Db 5101 AACCCCAACACAAACCCGACGACGACGACGACGACGACGACGACGACGACGACGAC 5160
QY 481 ACCCAATATGATTTTCACTTGCAGAGTGTTAACCTTGACCCGACGACGACGACGAC 540
Db 5161 ACCCAATATGATTTTCACTTGCAGAGTGTTAACCTTGACCCGACGACGACGACGAC 5220
QY 541 CAATCAACCTGCTGGGTATCTGCAAAAAGATTCGCAAAACAAAAACGAGAAAGAAC 600
Db 5221 CAATCAACCTGCTGGGTATCTGCAAAAAGATTCGCAAAACAAAAACGAGAAAGAAC 5280
QY 601 CACCAACAGGCTCAAAAAACCACTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660
Db 5281 CACCAACAGGCTCAAAAAACCACTTCAAGACCAACCAAAAAAGATCTCAAACTCA 5340
QY 661 AACCACTTAACCAAGAGTACCCACGACCAAGCCCAAGAGACCAACCATCAACAC 720
```

Db	5341	AACCACATAATCAAAGGAAGTACCACACCAAGGCCAAGAGAGCCAAACCATCAACAC	5400
Qy	721	CACCAAAACAAACATCAACATCACTGCTCAACCAACAACACAGAGAAATCCAAACT	780
Db	5401	CACCAAAACAAACATCACTAATCACTACTACTCACTCCACACACACAGGAAATCCAGAACT	5460
Qy	781	CACAAGTCAATGGAAACCTTCCTCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCA	840
Db	5461	CACAAGTCAATGGAAACCTTCCTCACTCAACCTCTCCGAAGGCAATCCAAAGCCCTTCA	5520
Qy	841	AGCTCCACAACATCGAGACACCCATCACACCCCTATCTCCAGCCACACAAACAGGCA	900
Db	5521	AGCTCTCAACAACCGAAGTACCCATCACACCTTATCTCCACCCAAACACACAGGCCA	5580
Qy	901	GTAGTATATTAATAAAAAA 919	
Db	5581	GTAGTACTTAATAAACATA 5599	

Search completed: May 1, 2003, 11:59:35  
Job time : 1875.32 secs

us-09-462-816-1.rng

Human RSV G-proc  
Respiratory syncyt

Respiratory syncyt

Respiratory syncytot  
Human respiratory  
RSV isolate 18537

DNA encoding the L

Nucleotide sequence

NUCLEOLIDE sequence  
DNA encoding the L

RSV 1.evertant 2B33

Nucleotide sequence

Nucleotide sequence  
 CNA encoding the I

Encodes Streptococcus

Nuclide sequence

RSV subgroup A clo  
RSV subgroup A prote  
RSV subgroup A W1  
respiratory Syncyt  
RSV protein anti  
RNA encoding a G2N  
nicotinic acid sequenc

n; vaccine;

XX Ewasyshyn ME, Klein MH;  
 XX WPI: 1993-243222/30.  
 DR P-PSDB: AAR39286.  
 XX  
 PT Multimeric hybrid genes and their chimeric proteins - are  
 PT vaccines against multiple pathogenic infections e.g.  
 PT para-influenza virus and respiratory syncytial virus  
 XX  
 PS Claim 11: Figure 7A-7D: 80pp: English.  
 XX  
 CC A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AA045683, AA045684) and RSV F and  
 CC G proteins (AA045685, AA045686).  
 CC  
 SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;  
 XX  
 Query Match 100.0%; Score 920; DB 14; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-211;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGCACAAATGTCACAAAGACCAAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 TCCAAACATGTCACAAAGACCAAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 61 CACTCTCAATCTTTATTTATTCATATCATTCGCGCTTATATTAAGTTAAATCTTAACTGT 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 61 CACTCTCAATCTTTATTTATTCATATCATTCGCGCTTATATTAAGTTAAATCTTAACTGT 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 121 AGCACAATCATATATTCATCTGCGAATGATATTCATCTTATATTAAGC 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 121 AGCACAATCATATATTCATCTGCGAATGATATTCATCTTATATTAAGC 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 181 CATCATATTCATATGCTCGCAACCAAGTCACACTACACTGCAATCATACAGA 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 181 CATCATATTCATATGCTCGCAACCAAGTCACACTACACTGCAATCATACAGA 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 241 TGCACAAAGCAGATTAAGACACAAACCCACATACCTGACTGAGATCTGAGCTGG 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 241 TGCACAAAGCAGATTAAGACACAAACCCACATACCTGACTGAGATCTGAGCTGG 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 301 AATCAGCTTCTCCATCTGTCTGAATTCATACAAACACACACCATCTAGCTTCAAC 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 301 AATCAGCTTCTCCATCTGTCTGAATTCATACAAACACACACCATCTAGCTTCAAC 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACAAAC 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACAAAC 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 421 AACCCAAACCAACCCAGCAAGCCCACTCAAAAGCAAGCAAAACCAACCAACCA 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 421 AACCCAAACCAACCCAGCAAGCCCACTCAAAAGCAAGCAAAACCAACCAACCA 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 481 ACCCAATATGATTTTCACTTCAAGTGTGTAATCTTGAACCTGAGCATATGACGAA 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 481 ACCCAATATGATTTTCACTTCAAGTGTGTAATCTTGAACCTGAGCATATGACGAA 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 541 CAATTCACCTGTGTGGCTATCTGCAAAAGATATCAAAACCAAGCAAGCAAGCA 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 541 CAATTCACCTGTGTGGCTATCTGCAAAAGATATCAAAACCAAGCAAGCAAGCA 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 601 CACACCAAGCTGACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 601 CACACCAAGCTGACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 661 AACCACTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 661 AACCACTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 721 CACCAAAACAAATCATACACTACACTGCTACCAACAAACACAGGAATCCAAACT 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 721 CACCAAAACAAATCATACACTACACTGCTACCAACAAACACAGGAATCCAAACT 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 781 CACAACTCAAAATGAAACCTTCCACTCAACCTCTCCGAGGAATCTTACGCTTCTCA 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 781 CACAACTCAAAATGAAACCTTCCACTCAACCTCTCCGAGGAATCTTACGCTTCTCA 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 841 AGCTCTCAACATCTGACACCCATCAACCCATCTGACCCACACACAGCA 900  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 841 AGCTCTCAACATCTGACACCCATCAACCCATCTGACCCACACACAGCA 900  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 901 GTAGTTATTTAAAAAAA 920  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 901 GTAGTTATTTAAAAAAA 920  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 2  
 AAX08421  
 ID AAX08421 standard; cDNA; 920 BP.  
 XX  
 AC AAX08421;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein gene of respiratory syncytial virus.  
 XX  
 XX G protein: respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator; ss.  
 XX  
 OS Respiratory syncytial virus (RSV).  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 8..904  
 FT /\*tag- a  
 FT /product- "Membrane bound G protein"  
 XX  
 W09904010-A1.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PE 16-JUL-1998; 98WO-CA00697.  
 XX  
 PR 18-JUL-1997; 97US-0896442.  
 XX  
 PA (CONN-) CONNAGHT LAB LTD.  
 XX  
 PI Klein MH, LI X, Sambhara S;  
 XX  
 DR WPI: 1999-132254/11.  
 DR P-PSDB: AAM96313.  
 XX  
 PT Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 3; Figure 2; 67pp: English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G protein's  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,

CC by usual immunisation and cell fusion methods.

XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 920; DB 20; Length 920;

Best Local Similarity 100.0%; Pred. No. 3.2e-211;

Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCGCAACATGTCCAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACTGGGA 60
DB 1 TCGCAACATGTCCAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACTGGGA 60
QY 61 CACTCTCAATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
DB 61 CACTCTCAATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCAAAATTCATTTATTCATTTCTGGCAATGATTAATTCACCTTCACTTAATTAAGC 180
DB 121 AGCAAAATTCATTTATTCATTTCTGGCAATGATTAATTCACCTTCACTTAATTAAGC 180
QY 181 CATCATTTATAGCTCGGCAACGACCAAGTCAACATACTGCAATCATACAGA 240
DB 181 CATCATTTATAGCTCGGCAACGACCAAGTCAACATACTGCAATCATACAGA 240
QY 241 TGCACACAGCCAGATTCACAGAACACACCCCAACATACCTCAGAGATCCTCAGCTTGG 300
DB 241 TGCACACAGCCAGATTCACAGAACACACCCCAACATACCTCAGAGATCCTCAGCTTGG 300
QY 301 AATGAGCTTTCGCAATGCTGTGTAATTCATTCACAAACCCGCACTACTAGCTTCAAC 360
DB 301 AATGAGCTTTCGCAATGCTGTGTAATTCATTCACAAACCCGCACTACTAGCTTCAAC 360
QY 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
DB 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
QY 421 AACCCAAACGACAGCCGCAAGCCCACTACAAACAAAGCCCAACCAACCAACCAACCAAC 480
DB 421 AACCCAAACGACAGCCGCAAGCCCACTACAAACAAAGCCCAACCAACCAACCAACCAAC 480
QY 481 ACCCAATATGATTTTCACTTCGAAAGTGTAACTTTGACCCGCGCATATGACACAA 540
DB 481 ACCCAATATGATTTTCACTTCGAAAGTGTAACTTTGACCCGCGCATATGACACAA 540
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAAC 600
DB 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAAC 600
QY 601 CACCACCAAGCTTACAAACCAACCTTCAAGACACCAACCAACCAACCAACCAACCTTCA 660
DB 601 CACCACCAAGCTTACAAACCAACCTTCAAGACACCAACCAACCAACCAACCAACCTTCA 660
QY 661 AACCACTAAACCAAGGAGTACCCACACCAAGCCCAAGAGACCCCAACCTTCAACAC 720
DB 661 AACCACTAAACCAAGGAGTACCCACACCAAGCCCAAGAGACCCCAACCTTCAACAC 720
QY 721 CACCAAAACCAATCACAACCTACCTGCTACCAACCAACCAACCAACCAACCAACCT 780
DB 721 CACCAAAACCAATCACAACCTACCTGCTACCAACCAACCAACCAACCAACCAACCT 780
QY 781 CACAGTCAATGGAACCTTCACATCAACCTTCCGAAGGCAATCTAAGCCCTTCTCA 840
DB 781 CACAGTCAATGGAACCTTCACATCAACCTTCCGAAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGCTTCACACATTCGAGACACCCATACAAACCTCATCTGCACCCCAACCAACCAAGC 900
DB 841 AGCTTCACACATTCGAGACACCCATACAAACCTCATCTGCACCCCAACCAACCAAGC 900
QY 901 GTAGTATTAAAAA 920
DB 901 GTAGTATTAAAAA 920

```

```

RESULT 3
ID AAN70784
XX AAN70784 standard; cDNA; 935 BP.
AC AAN70784;
XX
DT 05-APR-1991 (first entry)
XX
DE Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE G protein.
XX
KW Vaccine; ss.
XX
OS Human respiratory syncytial virus (HRSV).
XX
FH Key Location/Qualifiers
FT CDS 16..913
FT /tag= a
XX
PN W08704185-A.
XX
PD 16-JUL-1987.
XX
PE 23-DEC-1986; 86MO-US02756.
XX
PR 14-JAN-1986; 86US-0818740.
XX
PA (UYN-) UNIV OF N CAROLINA.
XX
PA (WERT/) WERTZ G W.
XX
PA (WERT/) WERTZ G W.
XX
DR WPI: 1987-206300/29.
DR P-PSDB; AAF70845.
XX
PT Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
XX
PS Disclosure; Chart 13; 57pp; English.
XX
CC A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women, or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
XX
SQ Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 93.2%; Score 857.6; DB 8; Length 935;
Best Local Similarity 95.8%; Pred. No. 3e-196;
Matches 881; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 TCGCAACATGTCCAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACTGGGA 60
DB 9 TCGCAACATGTCCAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACTGGGA 68
QY 61 CACTCTCAATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
DB 69 CACTCTCAATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAATCTGT 128
QY 121 AGCAAAATTCATTTATTCATTTCTGGCAATGATTAATTCACCTTCACTTAATTAAGC 180
DB 129 AGCAAAATTCATTTATTCATTTCTGGCAATGATTAATTCACCTTCACTTAATTAAGC 188
QY 181 CATCATTTATAGCTCGGCAACCAAGTCAACACTAAGCAATCATACAGA 240
DB 189 CATCATTTATAGCTCGGCAACCAAGTCAACACTAAGCAATCATACAGA 248
QY 241 TGCACACAGCCAGATTCACAGAACACACCCCAACATACCTCAGAGATCCTCAGCTTGG 300
DB 249 TGCACACAGCCAGATTCACAGAACACACCCCAACATACCTCAGAGATCCTCAGCTTGG 308
QY 301 AATGAGCTTTCGCAATGCTGTGTAATTCATTCACAAACCAACCAACCAATACTAGCTTCAAC 360

```

```

Db      309  AATCAGTCCCTCTAATCGCTGAAATTACACAAATTCACCCACATGACTGCTTCAAC 368
OY      361  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 420
Db      369  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 428
OY      421  AACCCACACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 480
Db      429  AACCTCAACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 488
OY      481  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 540
Db      489  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 548
OY      541  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 600
Db      549  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 608
OY      601  CACACCAAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATTCGAACCTCA 660
Db      609  CACTACCAAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATTCGAACCTCA 668
OY      661  AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAAGAGCCCAACCATCAACAC 720
Db      669  AACCACTAAATCAAGAGTACCCACCAACCAAGCCCAAGAGCCCAACCATCAACAC 728
OY      721  CACCAAAACCAACATCACAACCTACTGCTCAACCAACCAACCAAGGAAATCCAAACT 780
Db      729  CACCAAAACCAACATCACAACCTACTGCTCAACCAACCAACCAAGGAAATCCAAACT 788
OY      781  CACAAGTCAATGGAACCTTCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
Db      789  CACAAGTCAATGGAACCTTCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 848
OY      841  AGTCTCACACAACTCCGAGCAGCCATCACAACCTCATCTCCAAACCAACCAAGCCCA 900
Db      849  AGTCTCACACAACTCCGAGCAGCCATCACAACCTCATCTCCAAACCAACCAAGCCCA 908
OY      901  GTAGTTATTTAAAAAAA 920
Db      909  GTAGTTATTTAAAAAAA 928

RESULT 4
AAV18736
ID      AAV18736 .standard; cDNA; 935 BP.
XX      AC      AAV18736;
XX      DT      11-JUN-1998 (first entry)
XX      DE      HRSV glycoprotein G cDNA.
XX      KM      HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine; ss.
XX      OS      Human respiratory syncytial virus.
XX      FH      Key
FT      CDS      16..912
FT      FT      Location/Qualifiers
FT      FT      /tag= a
FT      FT      /product= glycoprotein_G
XX      XX      US5716823-A.
XX      PD      10-FEB-1998.
XX      PE      12-MAY-1997; 97US-0854783.
XX      PR      13-JUL-1988; 8805-0218737.
XX      PR      14-JAN-1986; 86US-0818740.
XX      PR      23-DEC-1986; 86MO-US02756.
XX      PR      11-JUN-1992; 92US-0897171.
```

```

PR      12-MAY-1997; 97US-0854783.
XX      XX      (PHAA ) PHARMACIA & UPJOHN CO.
XX      PI      Collins PL, Wertz GW;
XX      XX      WPI: 1998-144802/13.
XX      DR      P-PSDB; AAM47605.
XX      PT      Production of human respiratory syncytial virus glyco-protein F or G
XX      PT      - by culturing eukaryotic host cells transfected with corresponding
XX      PT      DNA
XX      PS      Example 1; Columns 27-28; 17pp; English.
XX      CC      The present sequence was used in the development of a novel method
XX      CC      for the production of human respiratory syncytial virus (HRSV)
XX      CC      glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
XX      CC      culturing eukaryotic host cells transfected with an isolated DNA
XX      CC      sequence encoding HRSV gpF or gpG. The gp can be used to prepare
XX      CC      vaccines against HRSV.
XX      SO      Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match      93.28; Score 857.6; DB 19; Length 935;
Best Local Similarity 95.88; Pred. No. 3e-196;
Matches 881; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY      1  TGCACACATGTCACAAAAACAGACCAACGACCGCTAAGCACTAGAAAAAGACTGGGA 60
Db      9  TGCACACATGTCACAAAAACAGACCAACGACCGCTAAGCACTAGAAAAAGACTGGGA 68
OY      61  CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATATAGTTAATCTGT 120
Db      69  CACTCTCAATCATTTATTTATTCATATCATGCGTCTTATATAGTTAATCTGTATCTGT 128
OY      121  AGCAAAATTCATATATCATTTCTGGCAATGATATATCAACTTCACTTAAATATACAGC 180
Db      129  AGCAAAATTCATATATCATTTCTGGCAATGATATATCAACTTCACTTAAATATGAGC 188
OY      181  CATCATATTCATAGGCTCGGCAACCAACAAAGTCACACTAACAAGTCAATACAGA 240
Db      189  CATCATATTCATAGGCTCGGCAACCAACAAAGTCACACTAACAAGTCAATACAGA 248
OY      241  TGCACAAAGCCAGATCAAGAACACACACCCCAACATACCTCACTGAGATCTCGCTTGG 300
Db      249  TGCACAAAGCCAGATCAAGAACACACACCCCAACATACCTCACTGAGATCTCGCTTGG 308
OY      301  AATCAGCTTCTCCATCTGTCTGAAATTCATCAACAACCAACCAACCATACTAGCTTCAAC 360
Db      309  AATCAGTCCCTCTAATCCGTCTGAAATTCATCAACAACCAACCAACCATACTAGCTTCAAC 368
OY      361  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 420
Db      369  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 428
OY      421  AACCCACACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 480
Db      429  AACCTCAACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 488
OY      481  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 540
Db      489  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 548
OY      541  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 600
Db      549  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 608
OY      601  CACACCAAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATTCGAACCTCA 660
Db      609  CACTACCAAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATTCGAACCTCA 668
OY      661  AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAAGAGCCCAACCATTAACAC 720
```

```

|||||
Db 669 AACCACTAAATCAAGAGAGTACCCACCAAGCCACAGAGAGCAACATCAACAC 728
QY 721 CACCAAAACAACTCAGCACTACACGTCGACCAACACAGCAAGAAATCCAAACT 780
Db 729 CACCAAAACAACTCAGCACTACACGTCGACCAACACAGCAAGAAATCCAGACT 788
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAGCCCTCTCA 840
Db 789 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAGCCCTCTCA 848
QY 841 AGTCTCACAACATCCGAGCAACCCATCACAACCCATCTCCACCCACCAACAGCGCA 900
Db 849 AGTCTCACAACATCCGAGCAACCCATCACAACCCATCTCCACCCACCAACAGCGCA 908
QY 901 GTAGTTATTAATAAAAAA 920
Db 909 GTAGTTACTTAATAAAAAA 928

```

## RESULT 5

AAQ29623 standard; DNA; 935 BP.

AAQ29623:

```

03-MAR-1993 (first entry)
HSRV glycoprotein G (gpG).
Vaccine: human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
major capsid protein; N: ss.
Human respiratory syncytial virus strain A2.

Key      Location/Qualifiers
FT CDS           16..912
FT             /tag= a
FT             /label= G-protein
FT             16..22
FT             /tag= b
FT             /note= "Oligonucleotide used to probe for
FT             full length CDNA"
FT             898..912
FT             /tag= C
FT             /note= "Oligonucleotide used to specifically
FT             prime the reverse transcription reaction
FT             for making the first strand of the CDNA"

```

```

US5149650-A.
22-SEP-1992.
14-JAN-1986; 86US-0818740.
14-JAN-1986; 86US-0818740.
PR 13-JUL-1988; 88US-0218737.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Collins PL, Wertz GW;
XX
DR WPI: 1992-340247/41.
XX
DR P-PSDB; AAR25302.
XX
PT Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
XX fragments
XX
PS Disclosure: Page 18; 21pp; English.
XX
CC The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid

```

CC protein N. The sequences and encoded proteins are useful for  
 CC preparing vaccines against HRSV. The vaccines can be used to confer  
 CC immunity against respiratory tract infections on human subjects.

SQ Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;

Query Match 93.0%; Score 856; DB 13; Length 935;

Best Local Similarity 95.7%; Pred. No. 7.3e-196;

Matches 880; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

QY 1 TCGAAGCATGTCCAAAACAAAGAGCAAGCGCCCTAAGACACTAGAAAAGCCTGGGA 60
Db 9 TCGAAGCATGTCCAAAACAAAGAGCAAGCGCCCTAAGACACTAGAAAAGCCTGGGA 68
QY 61 CACTCTCAATCATTTATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTGT 120
Db 69 CACTCTCAATCATTTATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTGT 128
QY 121 AGCACAATATCATTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 180
Db 129 AGCACAATATCATTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 188
QY 181 CATCATATTCATAGGCTCGGCAACCAAGTCACTAATCAATCAATCAATCAATCAAGA 240
Db 189 CATCATATTCATAGGCTCGGCAACCAAGTCACTAATCAATCAATCAATCAATCAAGA 248
QY 241 TGCACAAGCCGATCAAGACACAAACCCCAATACCTCACTAGATCTCACTCTTG 300
Db 249 TGCACAAGCCGATCAAGACACAAACCCCAATACCTCACTAGATCTCACTCTTG 308
QY 301 AATCAGCTTCTCCAAATCTGTGTGAATTTACATCAACCAACCAACCAATACAGCTCAAC 360
Db 309 AATCAGCTTCTCCAAATCTGTGTGAATTTACATCAACCAACCAACCAATACAGCTCAAC 368
QY 361 AACACCGAGATCAAGTCAAACTGTCAACCAACCAACCAACCAACCAACCAACCAAC 420
Db 369 AACACCGAGATCAAGTCAAACTGTCAACCAACCAACCAACCAACCAACCAACCAAC 428
QY 421 AACCCAAACACACCCAGCCAGCCAGCTACAAACCAACCAACCAACCAACCAACCAAC 480
Db 429 AACCCAAACACACCCAGCCAGCCAGCTACAAACCAACCAACCAACCAACCAACCAAC 488
QY 481 ACCCAATATGATTTTCACTTGAAGTGTTTTAACTTGTACCTGACGATATGACGAA 540
Db 489 ACCCAATATGATTTTCACTTGAAGTGTTTTAACTTGTACCTGACGATATGACGAA 548
QY 541 CAATGCAACCTGCTGGGTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
Db 549 CAATGCAACCTGCTGGGTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 608
QY 601 CACCACCAAGCCTTACAAAACCAACCAACCTTACAGACCAACCAACCAACCAACCAAC 660
Db 609 CACTACCAAGCCTTACAAAACCAACCAACCTTACAGACCAACCAACCAACCAACCAAC 668
QY 661 AACCACTAACTAAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAAC 720
Db 669 AACCACTAACTAAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAAC 728
QY 721 CACCAAAACAACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 780
Db 729 CACCAAAACAACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 788
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAGCCCTCTCA 840
Db 789 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAGCCCTCTCA 848
QY 841 AGTCTCACAACATCCGAGCAACCCATCACAACCCATCTCCACCCACCAACAGCGCA 900
Db 849 AGTCTCACAACATCCGAGCAACCCATCACAACCCATCTCCACCCACCAACAGCGCA 908
QY 901 GTAGTTATTAATAAAAAA 920
Db 909 GTAGTTACTTAATAAAAAA 928

```

RESULT 6  
AAT78440  
ID AAT78440 standard; DNA; 15222 BP.  
XX AAT78440;  
XX  
XX  
XX 17-MAR-1998 (first entry)  
XX  
XX Human respiratory syncytial virus strain A2.  
XX  
XX  
XX Antisense oligonucleotide; respiratory syncytial virus;  
XX RSV; treatment; infection; inhibition; strain A2; ss.  
XX  
XX Human respiratory syncytial virus.  
XX  
XX MO9729757-A1.  
XX  
XX 21-AUG-1997.  
XX  
XX 14-FEB-1997; 97WO-US02531.  
XX  
XX 15-FEB-1996; 96US-0011725.  
XX  
XX (CLEV-) CLEVELAND CLINIC FOUND.  
XX (USSH ) US NAT INST OF HEALTH.  
XX  
XX Cirino NM, LJ G, Silverman RH, Torrence PF, Xiao W;  
XX WPI; 1997-424748/39.  
XX  
XX Polynucleotide containing sequence anti-sense to region of RSV -  
XX connected via a linker to an activator of RNaseL, used to treat RSV  
XX infections  
XX  
XX Disclosure: Pages 47-51; 89pp; English.  
XX  
XX The present sequence was used in the preparation of a novel  
XX polynucleotide, comprising an antisense oligonucleotide, with a  
XX hydroxy group at one end, that is complementary to 15-20 bases of  
XX the anti-genomic RNA strand of a respiratory syncytial virus (RSV),  
XX a linker attached to the OH-end of the antisense oligonucleotide and  
XX an oligonucleotide activator of RNaseL attached to the linker. The  
XX polynucleotide can be used to treat RSV infections, which can also  
XX be treated by administration of the antisense oligonucleotide, so as  
XX to form a complex with activated RNase L in vivo. The  
XX polynucleotide can be transported across the cell membranes without  
XX carriers or permeability agents, and once introduced destroys  
XX antisense target RNA. It also inhibits RSV infection in vitro in a  
XX superior manner to the conventional drug, ribavirin.  
XX  
XX Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;  
XX  
XX  
XX Query Match 92.8%; Score 853.4; DB 18; Length 15222;  
XX Best Local Similarity 95.5%; Pred. No. 6.9e-195;  
XX Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
XX  
XX  
XX 1 TGCACATGTCGCAAAAGCAAGCAAGCAGCGCTTAAGACTGAAAGAGCTGGGA 60  
DB 4681 TGCACATGTCGCAAAAGCAAGCAAGCAGCGCTTAAGACTGAAAGAGCTGGGA 4740  
XX  
XX 61 CACTCTCAATCATTTATTTATTCATATCAGCGCTTATATTAATCTTAATCTGT 120  
DB 4741 CACTCTCAATCATTTATTTATTCATATCAGCGCTTATATTAATCTTAATCTGT 4800  
XX  
XX 121 AGCACAATACATTTATTCCTTCTGGCAATGATATCTCACTTCACTTATTAACAG 180  
DB 4801 AGCACAATACATTTATTCCTTCTGGCAATGATATCTCACTTCACTTATTAACAG 4860  
XX  
XX 181 CATCATTTATGAGCTGGGCAAGCAAGCTACACTTAACATGCAATCATTAAGA 240  
DB 4861 CATCATTTATGAGCTGGGCAAGCAAGCTACACTTAACATGCAATCATTAAGA 4920

XX 241 TGCACAAGCCAGATCAGAAACAAACCCAGATACCTTCACTAGAGATCCCTGAGCTGG 300  
DB 4921 TGCACAAGCCAGATCAGAAACAAACCCAGATACCTTCACTAGAGATCCCTGAGCTGG 4980  
XX  
XX 301 AATCAGCTCTCCATCTGCTGTGTGAAATTTACATGACAAACCCAGCAGCATACTAGCTTCAAC 360  
DB 4981 AATCAGCTCTCCATCTGCTGTGTGAAATTTACATGACAAACCCAGCAGCATACTAGCTTCAAC 5040  
XX  
XX 361 AACACCAAGAGTCAAGTCAAACTGCAACCCAGCAACAGTCAACACTTAAACCAACCAAC 420  
DB 5041 AACACCAAGAGTCAAGTCAAACTGCAACCCAGCAACAGTCAACACTTAAACCAACCAAC 5100  
XX  
XX 421 AACCCAAACAAACCCAGCAAGCCCACTTACAAAACAAAGCCAAACCAACCAACCAACAA 480  
DB 5101 AACCCAAACAAACCCAGCAAGCCCACTTACAAAACAAAGCCAAACCAACCAACCAACAA 5160  
XX  
XX 481 ACCCAATTAATGATTTTCACTTGTGAAGTGTGTTTAACTTTGTACCTGCAAGCATATGACGAA 540  
DB 5161 ACCCAATTAATGATTTTCACTTGTGAAGTGTGTTTAACTTTGTACCTGCAAGCATATGACGAA 5220  
XX  
XX 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATGCAAAACCAAAACCAAGAAAGAAAC 600  
DB 5221 CAATCCAACTGCTGGGCTATCTGCAAAAGATATGCAAAACCAAAACCAAGAAAGAAAC 5280  
XX  
XX 601 CACCACCAAGGCTTACAAAACCAACCTTCAAGACCAACCAAAACCAAGATCTCAACCTCA 660  
DB 5281 CACCACCAAGGCTTACAAAACCAACCTTCAAGACCAACCAAAACCAAGATCTCAACCTCA 5340  
XX  
XX 661 AACCACTTAACCAAGAGATGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720  
DB 5341 AACCACTTAACCAAGAGATGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 5400  
XX  
XX 721 CACCAAAACCAAGATGACAGATGACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 780  
DB 5401 CACCAAAACCAAGATGACAGATGACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 5460  
XX  
XX 781 CACAAGTCAAAATGGAACCTTCACTCACTCACTCTCTCGAAGGCAATTAAGCCCTTCTCA 840  
DB 5461 CACAAGTCAAAATGGAACCTTCACTCACTCACTCTCTCGAAGGCAATTAAGCCCTTCTCA 5520  
XX  
XX 841 AGTCGCCAACAATCCGCGACCCATGCAACAGCCGATCTGACCAACCAACCAACCAACGCA 900  
DB 5521 AGTCGCCAACAATCCGCGACCCATGCAACAGCCGATCTGACCAACCAACCAACCAACGCA 5580  
XX  
XX 901 GTAGTTATTTAAAAA 919  
DB 5581 GTAGTTATTTAAAAA 5599  
XX  
XX  
XX RESULT 7  
XX AAT63430  
XX ID AAT63430 standard; DNA; 15223 BP.  
XX  
XX AAT63430;  
XX  
XX 02-JUL-1997 (first entry)  
XX  
XX Respiratory syncytial virus anti-genome.  
XX  
XX RSV; vaccine; gene therapy; upper respiratory tract disease; ss.  
XX  
XX Human respiratory syncytial virus strain A2.  
XX  
XX WO9712032-A1.  
XX  
XX 03-APR-1997.  
XX  
XX 27-SEP-1996; 96WO-US15524.  
XX  
XX 27-SEP-1995; 95US-0007083.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.



XX Collins PL;  
 XX WPI; 1997-212893/19.  
 XX  
 PT Infectious respiratory syncytial virus particles - useful for  
 PT treatment of RSV or gene therapy of upper respiratory tract diseases  
 XX  
 PS Claim 46; Page 43-51; 66pp; English.  
 XX

CC A human respiratory syncytial virus (RSV) anti-genome sequence  
 CC (AATG3430) is the 5' to 3' positive-sense sequence of RSV; the genome  
 CC itself is negative-sense. It was synthesised in segments by RT-PCR  
 CC using intracellular RSV mRNA or genomic RNA isolated from purified  
 CC viruses as template. Restriction site markers were introduced by  
 CC incorporating the changes into the primers used for RT-PCR. The  
 CC recombinant sequence can be expressed with a nucleocapsid protein,  
 CC a nucleocapsid phosphoprotein, a large polymerase protein and an  
 CC RNA elongation factor to produce isolated infectious RSV particles  
 CC useful for generating vaccines against RSV. Recombinant RSV genome  
 CC or antigenome can also be used as a vector for gene therapy of the  
 CC upper respiratory tract.  
 CC  
 XX

Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

Query Match 92.8%; Score 853.4; DB 18; Length 15223;  
 Best Local Similarity 95.5%; Pred. No. 6.9e-195;

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCACATGTCGCAAAACAGAGCAGCAGCCGCTAGACACTGTGAAAGACCTGGGA 60  
 DB 4682 TGCACATGTCGCAAAACAGAGCAGCAGCCGCTAGACACTGTGAAAGACCTGGGA 4741  
 QY 61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATTAAGTTAAATCTTAACTCT 120  
 DB 4742 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATTAAGTTAAATCTTAACTCT 4801  
 QY 121 AGCACAATTCACATTTATTCATATCATGCGGCTTATTAAGTTAAATCTTAACTCT 180  
 DB 4802 AGCACAATTCACATTTATTCATATTCATGCGGCTTATTAAGTTAAATCTTAACTCT 4861  
 QY 181 CATCATATTCATATGCTCGGCAAAACAGAGCAGCAGCCGCTAGACACTGTGAAAGACCTGGGA 240  
 DB 4862 CATCATATTCATATGCTCGGCAAAACAGAGCAGCAGCCGCTAGACACTGTGAAAGACCTGGGA 4921  
 QY 241 TGCACACGCGAGATCAAGACACACCCCAATACCTACTCAGATCCTCAGCTTGG 300  
 DB 4922 TGCACACGCGAGATCAAGACACACCCCAATACCTACTCAGATCCTCAGCTTGG 4981  
 QY 301 AATCAGCTTCTCCATCTCTCTGAAATTCATCACAACCCAGCCATCTAGCTTCAAC 360  
 DB 4982 AATCAGCTTCTCCATCTCTCTGAAATTCATCACAACCCAGCCATCTAGCTTCAAC 5041  
 QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCAGCAGTCAAGCTTAAACACCAACCAAC 420  
 DB 5042 AACACGAGAGTCAAGTCAAACTGCAACCCAGCAGTCAAGCTTAAACACCAACCAAC 5101  
 QY 421 AACCCAAACACCAACCCAGCAGTCAAAACACCCCAACCAACCAACCAACCAACCAAC 480  
 DB 5102 AACCTCAAAACACCAACCCAGCAGTCAAAACACCCCAACCAACCAACCAACCAACCAAC 5161  
 QY 481 ACCCAATATGATTTTCACTTGTGAAGTGTTAACCTTGTACCTGACAGATATGACGCA 540  
 DB 5162 ACCCAATATGATTTTCACTTGTGAAGTGTTAACCTTGTACCTGACAGATATGACGCA 5221  
 QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACCAAGGAGGAGGAGGAGGAG 600  
 DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACCAAGGAGGAGGAGGAGGAG 5281  
 QY 601 CACGACCAAGCTACAAAAAACAACCTCTCAAGACCAACCAAAAAAGATCTCAAACTCA 660  
 DB 5282 CACTACCAAGGCGCAAAAAAACAACCTCTCAAGACCAACCAAAAAAGATCTCAAACTCA 5341

QY 661 AACCACTAAACCAAGAGATACCCACGACCAAGCCCAAGAGAGGACCAACATCAACAC 720  
 DB 5342 AACCACTAAATCAAGAGAGTACCCACGACCAAGCCCAAGAGAGGAGGAGGAGGAGGAG 5401  
 QY 721 CACCAAAACCAACATTCACATACACTGCTGCACCAACCAACCAACCAACCAACCAACCAAC 780  
 DB 5402 CACCAAAACCAACATTCACATACACTGCTGCACCAACCAACCAACCAACCAACCAACCAAC 5461  
 QY 781 CACAGTCAAAATGGAACCTTCCACTCAACCTCTCCCAAGGCAATCTTACCTTCTCA 840  
 DB 5462 CACAAAGTCAAAATGGAACCTTCCACTCAACCTCTCCCAAGGCAATCTTACCTTCTCA 5521  
 QY 841 ACTCTCACAACATCCGAGACCCATCAACACCCCTCATCTCCACCAACCAACCAACCAACCAAC 900  
 DB 5522 ACTCTCTCAACATCTCCGAGTACCCATCAACACCTTCACTCTCCACCAACCAACCAACCAAC 5581  
 QY 901 GTAGTTATTTAAAAA 919  
 DB 5582 GTAGTTACTTTAAAAACATA 5600

RESULT 8

AAV17553  
 ID AAV17553 standard; cDNA; 15223 BP.

AAV17553;

20-JUL-1998 (first entry)

Respiratory syncytial virus antigenome.

RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.

Human respiratory syncytial virus p46.

WO9802530-A1.

22-JAN-1998.

15-JUL-1997; 97WO-US12269.

23-MAY-1997; 97US-0047634.

15-JUL-1996; 96US-0021773.

09-MAY-1997; 97US-0046141.

(USSH ) US DEPT HEALTH & HUMAN SERVICES.

Bukreyev AA, Collins PL, Juhász K, Murphy BR, Teng MN;  
 Whitehead SS;

WPI; 1998-110579/10.

Attenuated respiratory syncytial virus vaccines - useful to protect  
 individuals against RSV infection

Example 7; Page 188-195; 238pp; English.

This is the 5'-3' positive sense nucleotide sequence of  
 respiratory syncytial virus (RSV) p46. The genome is  
 negative-sense; the complete nucleotide sequence of the  
 wild-type B-1 virus has also been determined (see AAV17552).  
 A novel infectious recombinant RSV comprises a RSV genome or  
 antigenome, a major nucleocapsid (N) protein, a nucleocapsid  
 phosphoprotein (P), a large polymerase protein (L), and a RNA  
 polymerase elongation factor, where the recombinant RSV has at  
 least two attenuating mutations; one of the mutations specifying a  
 temperature-sensitive (ts) substitution at amino acid Phe521,  
 Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 nucleotide substitution in the gene-start sequence of gene M2.  
 Also claimed are: (1) an isolated infectious RSV particle which  
 comprises a recombinant RSV (antigenome, N, P, and L proteins, a  
 RNA polymerase elongation factor, where the (anti)genome is modified:  
 (1) to ablate or modulate expression of a SH, NS1, NS2 or G gene or

CC a cis-acting regulatory sequence; and (ii) by a termination codon  
 CC introduced within a selected gene, or by a change in sequence,  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),  
 CC cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/2B5 (ATCC VR 2342) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.

XX  
 XX  
 XX Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

Query Match 92.8%; Score 853.4; DB 19; Length 15223;  
 Best Local Similarity 95.5%; Pred. No. 6,9e-195;  
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAGACATGTCGCAAAAGAGGACGACGCGTAGACACTGAGAAAGACGCTGGGA 60  
 DB 4682 TGCAGACATGTCGCAAAAGAGGACGACGCGTAGACACTGAGAAAGACGCTGGGA 4741  
 QY 61 CACTGCAATCATTTATTTATTCATATCATGCGGCTTATATAGTTAAATCTGT 120  
 DB 4742 CACTGCAATCATTTATTTATTCATATCATGCGGCTTATATAGTTAAATCTGT 4801  
 QY 121 AGCACAATACATTTATTCATTTCTGCGAATGATTAATCTCACTTATTAATTACAGC 180  
 DB 4802 AGCACAATACATTTATTCATTTCTGCGAATGATTAATCTCACTTATTAATTACAGC 4861  
 QY 181 CATCATTTATGAGCTTGGGCAAAACGCAAGTCACACTCACTCACTCACTATACAGA 240  
 DB 4862 CATCATTTATGAGCTTGGGCAAAACGCAAGTCACACTCACTCACTCACTATACAGA 4921  
 QY 241 TGCACAAGCCAGATCAAGACCAACCCGACATACCTTCACTAGATCTGAGCTTGG 300  
 DB 4922 TGCACAAGCCAGATCAAGACCAACCCGACATACCTTCACTAGATCTGAGCTTGG 4981  
 QY 301 AATCAGTCTGCAATGCTGTGTAATTAATCATGCAAAACGACACATATAGCTTCAC 360  
 DB 4982 AATCAGTCTGCAATGCTGTGTAATTAATCATGCAAAACGACACATATAGCTTCAC 5041  
 QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCGACAGTCAAGTCAAGTCAAGTCAAGTCA 420  
 DB 5042 AACACGAGAGTCAAGTCAAACTGCAACCCGACAGTCAAGTCAAGTCAAGTCAAGTCA 5101  
 QY 421 AACCCAACACACCCGACGCAAGCCACTACAAAACGCAAAACGCAAAACGCAAAACGCA 480  
 DB 5102 AACCCAACACACCCGACGCAAGCCACTACAAAACGCAAAACGCAAAACGCAAAACGCA 5161  
 QY 481 ACCCAATTAAGATTTTACCTGCAAGTGTGTAATTTGTAACCTGCGAGCATATGACGCA 540  
 DB 5162 ACCCAATTAAGATTTTACCTGCAAGTGTGTAATTTGTAACCTGCGAGCATATGACGCA 5221  
 QY 541 CAATCAACCTGCTGGGCTATCTGCAAAAGATTAACCAAAACGAGGAAAGAAAC 600  
 DB 5222 CAATCAACCTGCTGGGCTATCTGCAAAAGATTAACCAAAACGAGGAAAGAAAC 5281  
 QY 601 CACCAACGAGCTTCAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 660  
 DB 5282 CACCAACGAGCTTCAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 5341  
 QY 661 AACCACTAAACCAAGGAGTACCCACCAAGCCGACAGAGGAGCAACATCAACAC 720  
 DB 5342 AACCACTAAACCAAGGAGTACCCACCAAGCCGACAGAGGAGCAACATCAACAC 5401  
 QY 721 CACCAAAACCAACATCAACATCACTGCTCACCACACACACAGAGAAATCCAAACT 780  
 DB 5402 CACCAAAACCAACATCAACATCACTGCTCACCACACACACAGAGAAATCCAAACT 5461  
 QY 781 CACAAGTCMAATGGAACCTTCCACTCAACCTCTCGAAGGCAATTAAGCCCTTCTCA 840

DB 5462 CACAAGTCMAATGGAACCTTCCACTCAACCTCTCGAAGGCAATTCACAGCCCTTCTCA 5521  
 QY 841 AGTCTCAACCAATCCGACGACCCATCAACACCTCATCTTCACCCACCAACACAGCCCA 900  
 DB 5522 AGTCTCTCAACCAATCCGAGTCCATCAACACCTTCATCTTCACCCACCAACAGCCCA 5581  
 QY 901 GTAGTTATTAATAAAAAA 919  
 DB 5582 GTAGTTACTTAATAAACATA 5600

#### RESULT 9

AAA88743 standard; cDNA: 15223 BP.

AC AAA88743;  
 DT 19-FEB-2001 (first entry)  
 XX  
 XX Respiratory syncytial virus D46 5'-3' positive sense sequence.

OS RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.  
 OS Chimeric - Human respiratory syncytial virus.  
 OS Chimeric - Bacteriophage T7.  
 XX Synthetic.

PH Key Location/Qualifiers  
 FT variation replace(4,6)  
 FT /tag- a  
 FT mutation 1099

FT /tag- b  
 FT /note= "single C insertion to create AflII site  
 in the NS2-N intergenic region"  
 FT mutation replace(1139,A)  
 FT /tag- c  
 FT /note= "creates NcoI site in N gene nontranslated  
 region"  
 FT mutation replace(1140,G)  
 FT /tag- d  
 FT /note= "creates NcoI site in N gene nontranslated  
 region"

FT mutation replace(5612,A)  
 FT /tag- e  
 FT /note= "creates StuI site in G/F intergenic region"  
 FT mutation replace(5616,A)  
 FT /tag- f  
 FT /note= "creates StuI site in G/F intergenic region"  
 FT mutation replace(7560,A)  
 FT /tag- g  
 FT /note= "creates SphI site in F/M2 intergenic region"

WO20061611-A2.

19-OCT-2000.

31-MAR-2000; 2000WO-US08802.

13-APR-1999; 99US-0291894.

(USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PI Collins PL, Murphy BR, Whitehead SS;  
 DR PI; 2000-679462/66.

PT Infectious chimeric respiratory syncytial virus (RSV) produced from  
 PT cloned nucleotide sequences, useful as a vaccine against diseases  
 PT caused by the virus, such as pneumonia and bronchiolitis -  
 XX Example 7; Page 262-268; 280pp; English.

XX The present sequence is that of the 5' to 3' positive-sense

QY	1	TGCAACATGTCCAAAAACAAGACCAACGACCGCTAAGACACTAGAAAAAGCTGGGA	60
Db	4682	TGCAAAACATGTCCAAAAACAAGAGACCAACGACCGCTAAGACATTAGAAAGACCTGGGA	4741
QY	61	CACCTCAATCATTTATATTATCATATCAGCGGCTATATTAAGTTAAATCTTAAATCTGT	120
Db	4742	CACCTCAATCATTTATTTATTTATTCATATCATCTGCTTATATTAAGTTAAATCTTAAATCTGT	4801
QY	121	AGCAACAATCACATTTATTCATTTCTGGCAATGTAAATCTCAACTTATTAATATACAGC	180
Db	4802	AGCAACAATCACATTTATTCATTTCTGGCAATGTAAATCTCAACTTATTAATATTCAGC	4861
QY	181	CATCATTTTCATATGGCTCGGCAAAACCAACAAGTCACTACACACTGCATATCATAGA	240
Db	4862	CATCATTTTCATATAGCTCGGCAAAACCAACAAGTCACTACACACTGCATATCATAGA	4921
QY	241	TGCACAAGCCAGCATACAGAACACAACCCCACTACCTCACTCGATGCTCGAGCTTGG	300
Db	4922	TGCACAAGCCAGCATCAAGAACACAACCCCACTATCTCAACCCGAATCTCTCAGCTTGG	4981
QY	301	AATCAGCTTTCATATCTGTCTGAAATTATCATACAAACACCAACATAGTTCCTAAC	360
Db	4982	AATCAGTCCCTTAAATCCGCTGTGAATTATCATACAAATTCACACATAGTTCCTAAC	5041
QY	361	AATCAGCAGAGTAAATGTCAAAACCTGCAACCCCAACAGTCAAGATATAAAACACACAAAC	420
Db	5042	AATCAGCAGAGTAAATGTCAAAACCTGCAATCCATGCATACCAAGTCAAGACCAAAACACACAAAC	5101
QY	421	AATCCAAACACAACCCAGCAGAGCCCACTACAAAAACAGGCCCAAAACCAACCCACCAACAA	480
Db	5102	AATCCAAACACAACCCAGCAGAGCCCAACCAAAACAGGCCCAAAACCAACCCACCAAGAA	5161
QY	481	AATCCAAATATGATTTTCATCTTGGAAGTGTAACTTTGATACCTTGACGATATGACAGAA	540
Db	5162	AATCCAAATATGATTTTCATCTTGGAAGTGTCAACTTTGATACCTTGACGATATGACAGAA	5221
QY	541	CAATCCAACTGTGTGGGTATCTGCAAAAAGATATACCAACCAAAACCCAGGAAGAAAC	600
Db	5222	CAATCCAACTGTGTGGGTATCTGCAAAAAGATATACCAACCAAAACCCAGGAAGAAAC	5281
QY	601	CATCACAAGCCTACAAAAAAACAACCTTCAAGACAACAAAAAAGATCTCAAACTCA	660
Db	5282	CATCACAAGCCTACAAAAAAACAACCTTCAAGACAACAAAAAAGATCTCAAACTCA	5341

QY 1 TGCACATGTCCAAAAACGACCAACCGCACCCTAGACACTAGAAAGACCTGGA 60



```
Db 61 AATCATTTATATTCATATCATATGCTGCTATATATTAATTAATCTTAATCTGTAGACAAA 120
Qy 128 ATCATATATATCATATTTGGCAATGATATATCAATTCACCTTATATATACAGCATATA 187
Db 121 ATCATATATATCATATTTGGCAATGATATATCAATTCACCTTATATATATGAGCCATATA 180
Qy 188 TTCTACCTTCGGCAAAACCAAAAGTGCACACTTAACCACTCAATCATATACAGATGCACA 247
Db 181 TTCTACCTTCGGCAAAACCAAAAGTGCACACCAACCACTCAATCATATACAGATGCACA 240
Qy 248 AGCCAGATCAAGAACACAAACCCCAACATACCTCACTAGATGCTCACTGTGGAAATCAGC 307
Db 241 AGCCAGATCAAGAACACAAACCCCAACATACCTCACTAGATGCTCACTGTGGAAATCAGT 300
Qy 308 TTCTCAATCTGTGTAATTTATACATCAACCAACCACTTACCTAGTGTGTTAAACACCA 367
Db 301 CCCTTATATCGCTGTGAATTTATACATCAACCAACCACTTACCTAGTGTGTTAAACACCA 360
Qy 368 GGAATCAAGTCAAACTGCACACCCACACAGTCAAGACTAAAAACACACACACACCA 427
Db 361 GGAATCAAGTCAAACTGCACACCCACACAGTCAAGACTAAAAACACACACACCA 420
Qy 428 ACACAAACCCAGCAAGCCCACTACAAAACACAGCCCAAAACCAACCAACCAACCAAT 487
Db 421 ACACAAACCCAGCAAGCCCACTACAAAACACAGCCCAAAACCAACCAACCAACCAAT 480
Qy 488 AATGATTTTCACTTCGAAAGTGTAACTTGTACCTGACGATATGACACACCAATCCA 547
Db 481 AATGATTTTCACTTCGAAAGTGTAACTTGTACCTGACGATATGACACCAATCCA 540
Qy 548 ACCTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAACCAACCAAC 607
Db 541 ACCTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAACCAACCAAC 600
Qy 608 AAGCTTCAAAAAAACCACTTCAAGACACCAAAAAAAGATCTCAAACTCAAACTCA 667
Db 601 AAGCTTCAAAAAAACCACTTCAAGACACCAAAAAAAGATCTCAAACTCAAACTCA 660
Qy 668 AAACCAAGAGAGTACCCACACCAAGCCACAGAGAGCCCAACCATCAACACACCAAA 727
Db 661 AAATCAAGAGAGTACCCACACCAAGCCACAGAGAGCCCAACCATCAACACACCAAA 720
Qy 728 ACAACATATCACTACACTGCTCACCACACACACACACAGAAATCCAAATCTCACAAGT 787
Db 721 ACAACATATCACTACACTGCTCACCACACACACACACAGAAATCCAAAGTCAACAAGT 780
Qy 788 CAATATGAAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847
Db 781 CAATATGAAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCT 840
Qy 848 ACAACATCGGAGCAGCCATCAACACCTCATCTCCACCAACACACACAGCCAGTAG 904
Db 841 ACAACATCGGAGTACCCATCAACACCTCATCTCCACCAACACACAGCCAGTAG 897

RESULT 12
AA59703/c
ID AA59703 standard; DNA; 15210 BP.
XX
AC AA59703;
XX
DT 26-JUL-1999 (first entry)
XX
DE Polynucleotide sequence of RSV strain A2.
XX
KW Antisense oligonucleotide; negative-strand RNA virus; activator; RNase L;
XX respiratory syncytial virus; RSV; influenza; mumps; rabies; ss.
OS Respiratory syncytial virus.
XX
XX WO9922742-A1.
XX
XX 14-MAY-1999.
PD
```

```
XX
XX 02-NOV-1998; 98WO-US23391.
PF
PR 03-NOV-1997; 97US-0962690.
XX
PA (CLEV-) CLEVELAND CLINIC FOUND.
PA (USSH ) US NAT INST OF HEALTH.
XX
PI Cirino NM, Li G, Player MR, Silverman RH, Torrence P;
PI Xiao W;
XX
XX WPI; 1999-326917/27.
PT New composition useful for inhibiting or treating infections against
PT negative-strand RNA virus
XX
PS Disclosure; Fig 1; 98p; English.
XX
CC The specification describes a composition comprising a polynucleotide
CC consisting of an antisense oligonucleotide containing a hydroxy group,
CC complementary to the genomic or antigenomic strand of a negative-strand
CC RNA virus; and an activator of RNase L. The polynucleotide is used to
CC inhibit, or treat, infection by negative-strand RNA viruses, specifically
CC respiratory syncytial virus (RSV) but also (para)influenza, mumps, and
CC rabies. The polynucleotide can cross cell membranes without requiring
CC carriers or permeabilizing agents, and can selectively cleave the RNA
CC targeted by the oligonucleotide. The present sequence represents
CC the polynucleotide sequence of RSV strain A2.
XX
SQ Sequence 15210 BP; 4232 A; 2351 C; 2700 G; 5919 T; 8 other;
XX
Query Match 90.3%; Score 830.4; DB 20; Length 15210;
Best Local Similarity 94.6%; Pred. No. 2.3e-169;
Matches 869; Conservative 3; Mismatches 41; Indels 6; Gaps 1;
Qy 1 TSCAAACATGTCACAAAACCAAGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
Db 10530 TSCAAACATGTCACAAAACCAAGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 10471
Qy 61 CACTCTCAATCACTTTATATTCATATCATTCGCGCTTATATAGTTAAATCTTAAATCTGT 120
Db 10470 CACTCTCAATCACTTTATATTCATATCATTCGCGCTTATATAGTTAAATCTTAAATCTGT 10417
Qy 121 ACCACAAATACATATTCATATTCGCAATGATATATCTCAACTTCACTTAATATACAGC 180
Db 10416 ACCACAAATACATATTCATATTCGCAATGATATATCTCAACTTCACTTAATATACAGC 10357
Qy 181 CATCATATTCATAGCTTGGCAACACCAAGTCACTACATAACACTGCAATCATACAGA 240
Db 10356 CATCATATTCATAGCTTGGCAACACCAAGTCACTACATAACACTGCAATCATACAGA 10297
Qy 241 TGCACAAAGCCAGATCAAGAACACAAACCCCAACACTCTCAAGATCTCTCAAGTCTGG 300
Db 10296 TGCACAAAGCCAGATCAAGAACACAAACCCCAACACTCTCAAGATCTCTCAAGTCTGG 10237
Qy 301 AATCAGCTTCTCCATCTGTCTGTAATTTATACATCAACAAACCAACCACTATAGCTTCAAC 360
Db 10236 AATCAGCTTCTCCATCTGTCTGTAATTTATACATCAACAAATCAACCACTATAGCTTCAAC 10177
Qy 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGCTAAGCTTAAAAACACACAC 420
Db 10176 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGCTAAGCTAAGCTTAAAAAC 10117
Qy 421 AACCAACACACACAGCCAGCCACAGTCAAAACCAAGCCCAAAACCAACCAACCAACCA 480
Db 10116 AACCTAAACACACAGCCAGCCACAGTCAAAACCAAGCCCAAAACCAACCAACCAACCA 10057
Qy 481 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTAACCTTGCAGCATATGACAGCA 540
Db 10056 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTAACCTTGCAGCATATGACAGCA 9997
Qy 541 CAATCAACCTGCTGGGCTATCTGCAAAAGATACCAACAAAAACCAAGAAAGAAAC 600
```

```
Db 9996 CAAATCCACCTGCTGGGCTATCTGCAAAAGAAATACCAAAACAAAACAGAGAAAC 9937
QY 601 CACCACCAAGCTTACAAAAAACCACTTCAAGACACCAAAAAAGATCTCAACCTCA 660
Db 9936 CACTACCACCCCAAAAAAACCAACCTCAAGACACCAAAAAAGATCTCAACCTCA 9877
QY 661 AACCACTAAACCAAGAAATACCCACCAAGACCAAGACCAAGACCAACATCAAC 720
Db 9876 AACCACTAAATCAAGAAATACCCACCAAGACCAAGACCAAGACCAACATCAAC 9817
QY 721 CACCAAAACCAACATCAACATCACTGCTCACCACCAACACCAAGAAATCAAACT 780
Db 9816 CACCAAAACCAACATCAACATCACTGCTCACCACCAACACCAAGAAATCAAACT 9757
QY 781 CACAAGCAAAATGAAACCTTCACTCAACCTCCGAAAGCAATCTAAGCCCTCTCA 840
Db 9756 CACAAGCAAAATGAAACCTTCACTCAACCTCCGAAAGCAATCTAAGCCCTCTCA 9697
QY 841 AGTCTCCACAACATCCGACGACCCATCACAAACCTCTCAACCCACCAACACAGCCA 900
Db 9696 AGTCTCCACAACATCCGACGACCCATCACAAACCTCTCAACCCACCAACACAGCCA 9637
QY 901 GTAGTATTAAAAAAA 919
Db 9636 GTAGTACTTAAAAACATA 9618
```

## RESULT 13

AAK08422 standard; cDNA; 715 BP.

AAK08422;

28-JUN-1999 (first entry)

G protein gene fragment of respiratory syncytial virus.

KM protein; respiratory syncytial virus; RSV; recombinant vector;

KM vaccine; immune response; immunogenicity; tPA; antibody;

KM tissue plasminogen activator; ss.

Respiratory syncytial virus (RSV).

Key Location/Qualifiers

FT CDS 1..702

FT /product= "Secreted G protein"

W09904010-A1.

28-JAN-1999.

16-JUL-1998; 98WO-CA00697.

18-JUL-1997; 97US-0896442.

(CONN-) CONNAUGHT LAB LTD.

Klein MH, Li X, Sambhara S;

WPI; 1999-132254/11.

P-PSDB; AAW96314.

Immunogenic composition for generating antibodies against

respiratory syncytial virus - comprises non-replicating vector

containing the protein G sequence, useful in protective vaccines and

to raise antibodies for diagnosis

Claim 8; Figure 3; 67pp; English.

The respiratory syncytial virus (RSV) G protein can be used in

vaccines by inserting the G protein gene into a non-replicating

vector. The G protein is placed under the control of alternative

CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G proteins  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.

SQ Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;

Query Match 77.7%; Score 715; DB 20; Length 715;

Best Local Similarity 100.0%; Pred. No. 4.6e-162;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 CACAAAGTCACTACATCACTGCAATCTACACAGTGCACACACCGATCAACACCA 265

Db 1 CACAAAGTCACTACATCACTGCAATCTACACAGTGCACACACCGATCAACACCA 60

QY 266 ACCCAACATACCTCAGTACAGATCTCAGCTTGGATTCAGCTTCTCCTCAATCTGTGA 325

Db 61 ACCCAACATACCTCAGTACAGATCTCAGCTTGGATTCAGCTTCTCCTCAATCTGTGA 120

QY 326 ATTACATCAAAACCCACCACTACTAGCTTCAACACACGAGATCAAGTCAACCTG 385

Db 121 ATTACATCAAAACCCACCACTACTAGCTTCAACACACGAGATCAAGTCAACCTG 180

QY 386 CAACCCACCAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 445

Db 181 CAACCCACCAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 240

QY 446 ACTACAAACCAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 505

Db 241 ACTACAAACCAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 300

QY 506 GTGTTAACTTGTACCTGACATATGACACCAACATCAACATCAACATCAACATCAAC 565

Db 301 GTGTTAACTTGTACCTGACATATGACACCAACATCAACATCAACATCAACATCAAC 360

QY 566 AAAAGATATCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

Db 361 AAAAGATATCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420

QY 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685

Db 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480

QY 686 ACCACCAAGCCCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745

Db 481 ACCACCAAGCCCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540

QY 746 CGCTTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805

Db 541 CGCTTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600

QY 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCAACATCAACATCAACATCA 865

Db 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCAACATCAACATCAACATCA 660

QY 866 TCACAAACCTCTCTTCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

Db 661 TCACAAACCTCTCTTCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715

RESULT 14

AAV38298 standard; DNA; 696 BP.

AAV38298;

26-OCT-1998 (first entry)

```
XX Respiratory syncytial virus glycoprotein G gene portion.
DE
XX
KM RSV; glycoprotein G; pneumococcal surface protein A; PspA;
KM infection; Streptococcus pneumoniae; sepsis; otitis media;
KM meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;
KM ss.
XX Respiratory syncytial virus.
XX
XX W09824927-A1.
XX
XX 11-JUN-1998.
XX
XX 04-DEC-1997; 97MO-US22847.
XX
XX 04-DEC-1996; 96US-0759505.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Curfiel DT, McDaniel LS;
XX
XX WPI: 1998-33343/29.
XX
XX Plasmid containing pneumococcal epitope for expression in eukaryotic
XX cells - useful for eliciting immunological response to pneumococcal
XX infection or sepsis
XX
XX Example 1; Fig 1C; 47pp; English.
XX
XX This is a portion of the respiratory syncytial virus glycoprotein G
XX (RSVG) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)
XX to create plasmid pG74. This plasmid contains a human
XX cytomegalovirus immediate early promoter and the RSVG gene portion
XX such that when an in-frame fusion is made, the resultant fusion
XX protein may be transported to, and anchored in, a mammalian cell
XX membrane where it can be exposed to the host immune system.
XX Insertion of pneumococcal surface protein A (PspA) coding sequence
XX created plasmid pKSD2601. Intramuscular immunisation of BALB/c
XX mice with pKSD2601 induced protection against an otherwise lethal
XX challenge with a capsular type 3 pneumococcus. A claimed plasmid
XX for expression of pneumococcal epitope DNA in eukaryotic cells
XX includes a promoter for driving expression in a eukaryotic cell
XX (e.g. HCMV-IE), DNA encoding a leader sequence (e.g. of RSVG)
XX and DNA encoding a pneumococcal epitope such as PspA. The invention
XX also provides a vaccine comprising the plasmid and a suitable
XX carrier or diluent, and optionally one or more cytokines or DNA
XX encoding them, or a bacterial delivery system. The vaccine is used
XX to elicit an immunological response in a host, including humans,
XX susceptible to pneumococcal infection or sepsis. The plasmid can
XX also be used to express a pneumococcal epitope of interest in vitro.*
XX
XX Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;
XX
XX Query Match 69.3%; Score 637.4; DB 19; Length 696;
XX Best Local Similarity 96.1%; Pred. No. 1.9e-143;
XX Matches 664; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 248 AGCCAGATATAACAAACAAACCCCAACATACCTCAGATGCTGACCTTGGAATCAGC 307
DB 241 AGCCAGATATCAAGAAACAAACCCCAACATACCTCAGATGCTGGAATCAGT 300
QY 308 TTCTCCAAATCTCTGGAATTAATACACAAACCCACCATATAGCTTCAAGTACACCA 367
DB 301 CCTCTAATCGCTGTGAATTAATACAAATACACCATATAGCTTCAAGTACACCA 360
QY 368 GGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGACTAAACACAAAC- AACCCA 426
DB 361 GGAGTCAAGTCAAAACCTGCAATCCCAACAGTCAAGACTAAACACAAAC- ACAAACCA 420
QY 427 AACACAAACCCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACCA- ACCCAA 486
DB 421 AACACAAACCCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACCA- ACCCAA 480
QY 487 TAATGATTTTCACTTGAAGTGTTAACCTTGTAACCTTGCAAGCATATGCAACCAATTC 546
DB 481 TAATGATTTTCACTTGAAGTGTTAACCTTGTAACCTTGCAAGCATATGCAACCAATTC 540
QY 547 AACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGAAAG- AACCAACCAC 606
DB 541 AACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGAAAG- AACCAACTAC 600
QY 607 CAAGGCTCAAAAAAACCACCTTCAAGACAACCAAAAAAATCTCAAACTCT- AACCAAC 666
DB 601 CAAGGCTCAAAAAAACCACCTTCAAGACAACCAAAAAAATCTCAAACTCT- AACCAACCCAC 660
QY 667 TAAACCAAGCAAGTACCCACCAACCAAGCC 697
DB 661 TAAATCAAGCAAGTACCCACCAACCAAGCC 691
XX
XX RESULT 15
XX AAV17552
XX ID AAV17552 standard; cDNA: 15225 BP.
XX
XX AAV17552;
XX
XX 20-JUL-1998 (first entry)
XX
XX Respiratory syncytial virus genome.
XX
XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.
XX
XX Human respiratory syncytial virus B-1.
XX
XX W09802530-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97MO-US12269.
XX
XX 23-MAY-1997; 97US-0047634.
XX
XX 15-JUL-1996; 96US-0021773.
XX
XX 09-MAY-1997; 97US-0046141.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;
XX Whitehead SS;
XX
XX WPI: 1998-110579/10.
XX
XX Attenuated respiratory syncytial virus vaccines - useful to protect
XX individuals against RSV infection
XX
XX Example 4; Page 195-202; 238pp; English.
XX
XX This is the complete nucleotide sequence of the wild-type B-1
XX respiratory syncytial virus (RSV). The genome is negative-sense;
XX the 5'-3' positive-sense sequence of D46 is provided in AAV17553.
```





GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:49:02 ; Search time 38.8257 Seconds  
(without alignments)  
7266.906 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacaacatgcccaaaaac.....gtagtataaaaaaaaaa 920

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued\_Patents\_NA.\*  
2: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCrus.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920	100.0	920	2	US-08-467-963C-7
2	920	100.0	920	2	US-08-838-189D-7
3	920	100.0	920	3	US-08-852-344D-7
4	920	100.0	920	3	US-08-344-639E-7
5	920	100.0	920	4	US-08-467-969A-7
6	920	100.0	920	4	US-08-467-961A-7
7	920	100.0	920	4	US-08-001-554A-7
8	894	97.2	894	2	US-08-467-963C-28
9	894	97.2	894	2	US-08-838-189D-28
10	894	97.2	894	2	US-08-852-344D-28
11	894	97.2	894	3	US-08-344-639E-28
12	853.4	92.8	15222	2	US-08-801-898A-23
13	853.4	92.8	15222	4	US-08-962-690-12
14	853.4	92.8	15223	2	US-08-892-403A-1
15	853.4	92.8	15223	4	US-08-720-132-1
16	441.4	48.0	15225	2	US-08-892-403A-2
17	212.8	28.1	1050	3	US-08-836-501-76
18	210.8	22.9	303	3	US-08-721-979A-1
19	210.8	22.9	303	3	US-08-836-504A-1
20	210.8	22.9	303	3	US-08-836-501-1
21	210.8	22.9	303	3	US-09-654-289-1
22	209.6	22.6	1071	3	US-08-836-501-77
23	207.6	22.6	303	3	US-08-721-979A-3
24	207.6	22.6	303	3	US-08-836-504A-3
25	207.6	22.6	303	4	US-08-836-501-3
26	207.6	22.6	303	4	US-09-654-289-3
27	201.2	21.9	303	3	US-08-721-979A-14

28	201.2	21.9	303	3	US-08-836-501-14	Sequence 14, Appl
29	201.2	21.9	303	4	US-09-654-289-14	Sequence 14, Appl
30	137	14.9	840	5	PCT-US91-08177-12	Sequence 12, Appl
31	137	14.9	7323	5	PCT-US91-08177-1	Sequence 1, Appl
32	128.6	14.0	183	3	US-08-721-979A-23	Sequence 23, Appl
33	128.6	14.0	183	3	US-08-836-501-23	Sequence 23, Appl
34	128.6	14.0	183	4	US-09-654-289-23	Sequence 23, Appl
35	126	13.7	303	3	US-08-721-979A-2	Sequence 2, Appl
36	126	13.7	303	3	US-08-836-504A-2	Sequence 2, Appl
37	126	13.7	303	4	US-09-654-289-2	Sequence 2, Appl
38	126	13.7	303	4	US-08-721-979A-30	Sequence 30, Appl
39	125.4	13.6	183	3	US-08-836-501-30	Sequence 30, Appl
40	125.4	13.6	183	4	US-09-654-289-30	Sequence 30, Appl
41	124.2	13.5	177	3	US-08-721-979A-24	Sequence 24, Appl
42	124.2	13.5	177	4	US-08-836-501-24	Sequence 24, Appl
43	124.2	13.5	177	4	US-09-654-289-24	Sequence 24, Appl
44	122.8	13.3	303	3	US-08-721-979A-4	Sequence 4, Appl
45	122.8	13.3	303	3	US-08-721-979A-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-467-963C-7  
Sequence 7, Application US/08467963C  
Patent No. 5968776  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
APPLICANT: EMASYSYN, Mary E  
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
City: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,963C  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/838,189  
FILING DATE: 16-Apr-1997  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-467-963C-7

Query Match	100.0%;	Score 920;	DB 2;	length 920;
Best Local Similarity	100.0%;	Pred. No. 4.5e-243;		
Matches 920;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	TGC	AAACATGTC	CAAA	AAACAG	AGCA	AGCAGCCGCT	AAC	ACTG	AA	AAACACCTGG	GA	60				
Db	1	TGC	AAACATGTC	CAAA	AAACAG	AGCA	AGCAGCCGCT	AAC	ACTG	AA	AAACACCTGG	GA	60				
QY	61	CAC	TCATCAT	ATTA	ATTCAT	ATCAT	TCATG	GGGCTT	ATAT	AA	GTAAATCTT	TAATCTGT	120				
Db	61	CAC	TCATCAT	ATTA	ATTCAT	ATCAT	TCATG	GGGCTT	ATAT	AA	GTAAATCTT	TAATCTGT	120				
QY	121	AGC	ACAATTC	ACAT	TTAT	TCAT	TTCTG	GC	AAATG	AT	TAATCT	CACTTA	TAATTAACGC	180			
Db	121	AGC	ACAATTC	ACAT	TTAT	TCAT	TTCTG	GC	AAATG	AT	TAATCT	CACTTA	TAATTAACGC	180			
QY	181	CAT	CATATTC	AT	AGC	TCG	GGC	AAAC	CAAC	AAAGT	CAC	ACTTA	CAAC	TCAT	CAAGA	240	
Db	181	CAT	CATATTC	AT	AGC	TCG	GGC	AAAC	CAAC	AAAGT	CAC	ACTTA	CAAC	TCAT	CAAGA	240	
QY	241	TGC	AAC	AGC	AGAT	C	AGAT	C	AGAT	C	AGAT	C	AGAT	C	AGAT	C	300
Db	241	TGC	AAC	AGC	AGAT	C	AGAT	C	AGAT	C	AGAT	C	AGAT	C	AGAT	C	300
QY	301	AAT	CAC	CTT	CT	CCAT	CT	CT	GT	GA	AA	TTAC	AT	CA	TCA	CA	360
Db	301	AAT	CAC	CTT	CT	CCAT	CT	CT	GT	GA	AA	TTAC	AT	CA	TCA	CA	360
QY	361	AAT	CAC	AG	AGT	C	AGT	C	AGT	C	AGT	C	AGT	C	AGT	C	420
Db	361	AAT	CAC	AG	AGT	C	AGT	C	AGT	C	AGT	C	AGT	C	AGT	C	420
QY	421	AAC	CCA	AG	CA	CA	ACC	AG	CC	CA	CT	AC	CA	AA	CA	AG	480
Db	421	AAC	CCA	AG	CA	CA	ACC	AG	CC	CA	CT	AC	CA	AA	CA	AG	480
QY	481	ACC	CA	AT	TA	T	GAT	TTT	CAC	TT	G	GA	GT	TTT	A	CTT	540
Db	481	ACC	CA	AT	TA	T	GAT	TTT	CAC	TT	G	GA	GT	TTT	A	CTT	540
QY	541	CA	AT	TC	CA	AC	CT	G	T	G	G	G	C	T	A	T	600
Db	541	CA	AT	TC	CA	AC	CT	G	T	G	G	G	C	T	A	T	600
QY	601	CAC	CA	CA	AG	CC	T	AC	AAAA	AA	CC	CA	CC	TT	CA	AG	660
Db	601	CAC	CA	CA	AG	CC	T	AC	AAAA	AA	CC	CA	CC	TT	CA	AG	660
QY	661	AAC	CA	CT	AA	AC	CA	AA	GA	AT	ACC	CAC	CA	AG	AT	ACC	720
Db	661	AAC	CA	CT	AA	AC	CA	AA	GA	AT	ACC	CAC	CA	AG	AT	ACC	720
QY	721	CAC	CA	AA	AA	CA	AA	CA	CT	CA	CT	CA	CT	CA	CT	CA	780
Db	721	CAC	CA	AA	AA	CA	AA	CA	CT	CA	CT	CA	CT	CA	CT	CA	780
QY	781	CAC	AG	TC	CA	AA	T	G	AA	CT	TC	CA	CT	CA	CT	CA	840
Db	781	CAC	AG	TC	CA	AA	T	G	AA	CT	TC	CA	CT	CA	CT	CA	840
QY	841	AG	TC	TC	CA	CA	CA	TC	CG	AG	CA	CC	AT	C	A	A	900
Db	841	AG	TC	TC	CA	CA	CA	TC	CG	AG	CA	CC	AT	C	A	A	900
QY	901	G	T	A	G	T	A	T	T	A	A	A	A	A	A	A	960
Db	901	G	T	A	G	T	A	T	T	A	A	A	A	A	A	A	960

RESULT 2  
US-08-838-189D-7  
; Sequence 7, Application US/08838189D

```

Patent No. 5998169
GENERAL INFORMATION:
APPLICANT: KLEIN, MICHEL H
APPLICANT: DU, Run-Pan
APPLICANT: EWAJSHTN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: SIM & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-838-189D-7
Query Match 100.0%; Score 920; DB 2; Length 920;
Best Local Similarity 100.0%; Pctd. No. 4, 5e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	TGCAAAACATGTCACAAAACAAAGGACCAACGCCACCGGTAAAGACATAGAAAAGACCTGGGA	60
Db	1	TGCAAAACATGTCACAAAACAAAGGACCAACGCCACCGGTAAAGACATAGAAAAGACCTGGGA	60
QY	61	CATCTCATCATCTTATATATTCATATCATGCGGCTATATAGTTAAATCTTAAATCTGT	120
Db	61	CATCTCATCATCTTATATATTCATATCATGCGGCTATATAGTTAAATCTTAAATCTGT	120
QY	121	AGCACAATCACATTTATTCATCTTGGCAATGATTAATCTCAACTCTATTAAATTACAGC	180
Db	121	AGCACAATCACATTTATTCATCTTGGCAATGATTAATCTCAACTCTATTAAATTACAGC	180
QY	181	CATCATATTCATAGCGTCGGCAAAACGCAAGTCACACTATACACTGCAATCATATACAGA	240
Db	181	CATCATATTCATAGCGTCGGCAAAACGCAAGTCACACTATACACTGCAATCATATACAGA	240
QY	241	TGCAACCAACGCGATGCAAGAACCAACCCCAACATACCTCTCAGATTCCTCAGCTTGG	300
Db	241	TGCAACCAACGCGATGCAAGAACCAACCCCAACATACCTCTCAGATTCCTCAGCTTGG	300
QY	301	AATCAGCTTCTCAATCTGTGTGAATTACATACAAACCCACAGCATATAGCTTCAC	360
Db	301	AATCAGCTTCTCAATCTGTGTGAATTACATACAAACCCACAGCATATAGCTTCAC	360

361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTAAACAAACACAC 420  
361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTAAACAAACACAC 420  
421 AACCCAAACACACCCAGCAGCCACTACAAACAAACGCGCAACAAACACCAACAA 480  
421 AACCCAAACACACCCAGCAGCCACTACAAACAAACGCGCAACAAACACCAACAA 480  
481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540  
481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540  
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACAAACAAACGAGAAAGAAAC 600  
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACAAACAAACGAGAAAGAAAC 600  
601 CACCAACAGGCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660  
601 CACCAACAGGCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660  
661 AACCACTAAACCAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720  
661 AACCACTAAACCAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720  
721 CACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
721 CACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
781 CACCAAGTCAATGGAACCTTCCATCAACCTCTCGAAGGCAATCTAAGCCCTTCA 840  
781 CACCAAGTCAATGGAACCTTCCATCAACCTCTCGAAGGCAATCTAAGCCCTTCA 840  
841 AGCTTCACACATCGGAGCAACCAATCAACCTCTCTGACCCCAACACACGCGCA 900  
841 AGCTTCACACATCGGAGCAACCAATCAACCTCTCTGACCCCAACACACGCGCA 900  
901 GTAGTATATTAATAAAAAA 920  
901 GTAGTATATTAATAAAAAA 920

RESULT 3  
US-08-852-344D-7  
Sequence 7, Application US/08852344D  
Patent No. 6017539  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michael H  
APPLICANT: DU, Run-Pan  
APPLICANT: EWASYSHV, Mary E  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,344D  
FILING DATE: 07-MAY-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/344,639  
FILING DATE: 14-NOV-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-852-344D-7  
Query Match 100.0%; Score 920; DB 3; Length 920;  
Best Local Similarity 100.0%; Pred. No. 4,5e-243;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 TGCACAAATGTCACAAAACAGAGCAACGACCGCTAGACACTGAAAGAGCTGGGA 60  
1 TGCACAAATGTCACAAAACAGAGCAACGACCGCTAGACACTGAAAGAGCTGGGA 60  
61 CACTCTCAATCATTTATTTATCATATCATGCGGCTTATATTAATCTTAATCTGT 120  
61 CACTCTCAATCATTTATTTATCATATCATGCGGCTTATATTAATCTTAATCTGT 120  
121 AGCACAATCATCATTTATTCATTTGCGCATATATATCTCAATTTAACTTACG 180  
121 AGCACAATCATCATTTATTCATTTGCGCATATATATCTCAATTTAACTTACG 180  
181 CATCATATTCATTCCTTCGCAACCAACCAAGTACACTAAGCTCAATCTATACAGA 240  
181 CATCATATTCATTCCTTCGCAACCAACCAAGTACACTAAGCTCAATCTATACAGA 240  
241 TGCACAAGCCAGATCAAGAACACACACCCACATACCTCAGTACGATCTTGG 300  
241 TGCACAAGCCAGATCAAGAACACACACCCACATACCTCAGTACGATCTTGG 300  
301 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACCAACCAACCTTACTAGTTAC 360  
301 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACCAACCAACCTTACTAGTTAC 360  
361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACCAACAC 420  
361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACCAACAC 420  
421 AACCCAAACACACCCAGCAGCCACTACAAACAAACGCGCAACAAACACCAACAA 480  
421 AACCCAAACACACCCAGCAGCCACTACAAACAAACGCGCAACAAACACCAACAA 480  
481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540  
481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540  
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACAAACAAACGAGAAAGAAAC 600  
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACAAACAAACGAGAAAGAAAC 600  
601 CACCAACAGGCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660  
601 CACCAACAGGCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660  
661 AACCACTAAACCAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720  
661 AACCACTAAACCAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720  
721 CACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
721 CACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780



Thu May 1 14:50:26 2003

us-09-462-816-1.rni

Page 5

```

CORRESPONDENCE ADDRESS:
ADDRESSEE:  Slim & McBurney
STREET:  330 University Avenue, 6th Floor
CITY:  Toronto
STATE:  Ontario
COUNTRY:  Canada
ZIP:  M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE:  Floppy disk
OPERATOR:  IBM PC compatible
SOFTWARE SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/08/467, 969A
FILING DATE:  06-JUN-1995
CLASSIFICATION:  435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  US 08/001,554
FILING DATE:  06-JAN-1993
CLASSIFICATION:  435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  GB 9200117.1
FILING DATE:  06-JAN-1992
CLASSIFICATION:  435
ATTORNEY/AGENT INFORMATION:
NAME:  Stewart, Michael I
REGISTRATION NUMBER:  24, 973
REFERENCE/DOCKET NUMBER:  1038-475 MIS:bn
TELECOMMUNICATION INFORMATION:
TELEPHONE:  (416) 595-1155
TELEFAX:  (416) 595-1163
TELEX:  065-24567 SIMBAS
INFORMATION FOR SEQ ID NO:  7:
SEQUENCE CHARACTERISTICS:
LENGTH:  920 base pairs
TYPE:  nucleic acid
STRANDEDNESS:  single
TOPOLOGY:  linear
MOLECULE TYPE:  DNA (genomic)
OS-08-467-969A-7

```

Query Match	100.0%;	Score 920;	DB 4;	Length 920;
Best Local Similarity	100.0%;	Pred. No. 4.5e-243;		
Matches 920;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	421	AATCCAAACACACACCCAGCAGACCCACTCTACAAACACACGCCAAAACAAACACACACAC	480
Db	421	AATCCAAACACACACCCAGCAGACCCACTCTACAAACACACGCCAAAACAAACACACACAC	480
QY	481	ACCCAAATATGATTTTCTCACTTCGAAAGTGTTAACCTTTGACCTCGAGCAATATCCAGCA	540
Db	481	ACCCAAATATGATTTTCTCACTTCGAAAGTGTTAACCTTTGACCTCGAGCAATATCCAGCA	540
QY	541	CAATCCAACTCTGCTGGCTATCTGC AAAAGATATCCAAACAAAAAACACAGGAAAAAAC	600
Db	541	CAATCCAACTCTGCTGGCTATCTGC AAAAGATATCCAAACAAAAAACACAGGAAAAAAC	600
QY	601	CACCAACCAAGCTACAAAAAAACCAACCTTCACACACAAACAAAAAAGATCTCAACCTCA	660
Db	601	CACCAACCAAGCTACAAAAAAACCAACCTTCACACACAAACAAAAAAGATCTCAACCTCA	660
QY	661	AACCACTAATACCAAAAGCAAGTACCCACCCACCAAGCCACAGAGAGCCAAACCATCAACAC	720
Db	661	AACCACTAATACCAAAAGCAAGTACCCACCCACCAAGCCACAGAGAGCCAAACCATCAACAC	720
QY	721	CACCAAAACAAACATCAACATCACTGTCTACCAACACACACACAGCAAGGAAATCCAAACT	780
Db	721	CACCAAAACAAACATCAACATCACTGTCTACCAACACACACACAGGAAATCCAAACT	780
QY	781	CACCAAGTCAATTGAACACTTTCACACTCAACTCTCTCCGAAGGCAATTAAGCTCTCTCA	840
Db	781	CACCAAGTCAATTGAACACTTTCACACTCAACTCTCTCCGAAGGCAATTAAGCTCTCTCA	840
QY	841	AGTCTCCACAACATTCGAGACCCATCACAAACCTATATTCACCCACACACAAACAGGCA	900
Db	841	AGTCTCCACAACATTCGAGACCCATCACAAACCTATATTCACCCACACACAAACAGGCA	900
QY	901	GTAGTTATTAATAAAAAAAAAA 920	
Db	901	GTAGTTATTAATAAAAAAAAAA 920	

RESULT 6  
US-08-467-961A-7  
Sequence 7, Application US/08467961A  
Patent No. 6171/83  
GENERAL INFORMATION:  
APPLICANT: Kieto, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: Ewaszyn, Mary E  
TITLE OF INVENTION: Chimeric Immunogens  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6TH FLOOR  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,961A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-476 MIS:bh  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 TELE: 065-24567 SIMBAS  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-467-961A-7

Query Match 100.0%; Score 920; DB 4; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-243;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 60  
 DB 1 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 60  
 QY 61 CACTCTCAATCATTTATTTATTCATTCATCGGGCTTAATTAAGTTAAATCTTGT 120  
 DB 61 CACTCTCAATCATTTATTTATTCATTCATCGGGCTTAATTAAGTTAAATCTTGT 120  
 QY 121 AGCACAATTCATTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 180  
 DB 121 AGCACAATTCATTTATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 180  
 QY 181 CATCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 240  
 DB 181 CATCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 240  
 QY 241 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 300  
 DB 241 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 300  
 QY 301 AATCAGCTTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCA 360  
 DB 301 AATCAGCTTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCA 360  
 QY 361 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 420  
 DB 361 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 420  
 QY 421 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 480  
 DB 421 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 480  
 QY 481 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 540  
 DB 481 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 540  
 QY 541 CAATTCAGCTTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 600  
 DB 541 CAATTCAGCTTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 600  
 QY 601 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 660  
 DB 601 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 660  
 QY 661 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 720  
 DB 661 AACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 720  
 QY 721 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 780  
 DB 721 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 780  
 QY 781 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 840  
 DB 781 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 840

DB 781 CACAGCAGAGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 840  
 QY 841 AGCTCCACAACTCCGAGACCCATCAGACACCTTATTCACCCCAACACAGCGCA 900  
 DB 841 AGCTCCACAACTCCGAGACCCATCAGACACCTTATTCACCCCAACACAGCGCA 900  
 QY 901 GTAGTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 920  
 DB 901 GTAGTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 920

## RESULT 7

US-08-001-554A-7  
 Sequence 7, Application US/08001554A  
 Patent No. 6225091

GENERAL INFORMATION:  
 APPLICANT: Klein, Michel H  
 APPLICANT: Du, Run-Pan  
 APPLICANT: Ewasyshtyn, Mary E  
 TITLE OF INVENTION: Chimeric Immunogens  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: 330 University Avenue, 6th Floor  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/001,554A  
 FILING DATE: 06-JAN-1993  
 CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-286  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-001-554A-7

Query Match 100.0%; Score 920; DB 4; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-243;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 60  
 DB 1 TGCAGCAATGTCAGCAAGGACCAAGCCGCTAAGACATAGAAAGACCTGGGA 60  
 QY 61 CACTCTCAATCATTTATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 120  
 DB 61 CACTCTCAATCATTTATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 120  
 QY 121 AGCACAATTCATTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 180  
 DB 121 AGCACAATTCATTTATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 180  
 QY 181 CATCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 240  
 DB 181 CATCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 240

QY	241	TGCAACAAGCCAGATGTCAGAAACACACACCCCAACATACCTCAGTACGAGATCCTCAGCTTGG	300
Db	241	TGCAACAAGCCAGATGTCAGAAACACACCCCAACATACCTCAGTACGAGATCCTCAGCTTGG	300
QY	301	AATCAGCTTCGCAATCTGTCTGAAATTAATGATGCAAAACCCAGCAGCATACTAGCTTCAAC	360
Db	301	AATCAGCTTCGCAATCTGTCTGAAATTAATGATGCAAAACCCAGCAGCATACTAGCTTCAAC	360
QY	361	AACACCGGAGACTCAAGTCAAAACCTGCAACCCACACACGATCAAGACTATAAACAACACAC	420
Db	361	AACACCGGAGACTCAAGTCAAAACCTGCAACCCACACACGATCAAGACTATAAACAACACAC	420
QY	421	AACCCAAACACACACCCGACGACAGCCCATACGAAAAACAAGCCAAACAAACACACCAACAA	480
Db	421	AACCCAAACACACACCCGACGACAGCCCATACGAAAAACAAGCCAAACAAACACACCAACAA	480
QY	481	ACCCAAATATATATTTTCACTTCGAAAGTGTTAACCTTTGTACCTGCGACATATGACGCAA	540
Db	481	ACCCAAATATATATTTTCACTTCGAAAGTGTTAACCTTTGTACCTGCGACATATGACGCAA	540
QY	541	CAATCCAAACCTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCAGAAAGAAAC	600
Db	541	CAATCCAAACCTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCAGAAAGAAAC	600
QY	601	CACCACCAAGCTCTACAAAAAACCCACACTTCTAAGACACACCAAAAAAGATCTCAACCTCA	660
Db	601	CACCACCAAGCTCTACAAAAAACCCACACTTCTAAGACACACCAAAAAAGATCTCAACCTCA	660
QY	661	AACCACTAAACCAAGGAGAGTACCACACCCACCAAGCCACGAAAGAGCAACCATCAACAC	720
Db	661	AACCACTAAACCAAGGAGAGTACCACACCCACCAAGCCACGAAAGAGCAACCATCAACAC	720
QY	721	CACCAAAAAACAACATCAACAATACACTGCTCACCACACACACCAAGGAATCCAAAACCT	780
Db	721	CACCAAAAAACAACATCAACAATACACTGCTCACCACACACCAAGGAATCCAAAACCT	780
QY	781	CACAAGTCAAAATGGAACCTTCCCACTAACCTCTCGAAGGCAATCTAAGCCCTTCTCA	840
Db	781	CACAAGTCAAAATGGAACCTTCCCACTAACCTCTCGAAGGCAATCTAAGCCCTTCTCA	840
QY	841	AGTCCCAACAACATCCGACGACCCCATCAACCCCTCATCTCCACCCCAACGACACAGCCCA	900
Db	841	AGTCCCAACAACATCCGACGACCCCATCAACCCCTCATCTCCACCCCAACGACACAGCCCA	900
QY	901	GTAGTATATTAATAAAAAA 920	
Db	901	GTAGTATATTAATAAAAAA 920	

RESULT 8  
US-08-467-963C-28  
Sequence 28, Application US/08467963C  
Patent No. 5968776  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
APPLICANT: EMASYSHYN, Mary E  
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONTERS PROTECTION AGAINST  
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

OY		8	ATGTCCAAAAAACAAGGACCAACGGACCGCTTAACACTCTGAAGAATAACCTGGGACACTTC	67
Dd		1	ATGTCCAAAAAACAAGGACCAACGGACCGCTTAACACTCTGAAGAATAACCTGGGACACTTC	60
OY		68	AATCATTTTATTATTTCATATCATATCAGGGCTTTTATTAAGTTAAATCTTAAATCTGTAGCAA	127
Dd		61	AATCATTTTATTATTTCATATCATATCAGGGCTTTTATTAAGTTAAATCTTAAATCTGTAGCAA	120
OY		128	ATTCACATTATTCATTCTTGCGCAATATATAATCCAATTCACCTATTATTAATTAACGCCATATA	187
Dd		121	ATTCACATTATTCATTCTTGCGCAATATATAATCCAATTCACCTATTATTAATTAAGCCATATA	180
OY		188	TTCATAGCCTGGCAAAACCAACAAGTACACTAACAACTGCATATATCAAGATGCAACA	247
Dd		181	TTCATAGCCTGGCAAAACCAACAAGTACACTAACAACTGCATATATCAAGATGCAACA	240
OY		248	AGCCAGATCAGACACACACCACCACTACCTCACTCAGAGTCTCAGCTTGGAATCAC	307
Dd		241	AGCCAGATCAGACACACACCACCACTACCTCACTCAGAGTCTCAGCTTGGAATCAC	300
OY		308	TTCTCCAAATCTGTGGAATTAACATTCACAACCAACCAACCATCTAGCTTCAACAACACA	367
Dd		301	TTCTCCAAATCTGTGGAATTAACATTCACAACCAACCAACCATCTAGCTTCAACAACACA	360
OY		368	GGAGTCAAATCAAACTTCACAACCCACAACAGTAAAGACTTAAAACACAACAACCCAA	427
Dd		361	GGAGTCAAATCAAACTTCACAACCCACAACAGTAAAGACTTAAAACACAACAACCCAA	420
OY		428	ACACAACCCAGAACCCCACTACAAACCAAGCCCAAAACCAACCAACCAACCCCAAT	487
Dd		421	ACACAACCCAGAACCCCACTACAAACCAAGCCCAAAACCAACCAACCAACCCCAAT	480
OY		488	AATGATTTTTCACTTGAAGTGTTTAACTTTGTACCTGGAGATATGACGACAACTCA	547
Dd		481	AATGATTTTTCACTTGAAGTGTTTAACTTTGTACCTGGAGATATGACGACAACTCA	540
OY		548	ACCTGCTGGCGAATCTGCAAAAAGATACCAACCAAAAACCGAGAAAGAAAACCAACACC	607
Dd		541	ACCTGCTGGCGAATCTGCAAAAAGATACCAACCAAAAACCGAGAAAGAAAACCAACACC	600
OY		608	AAGCCTACAAAAAAACAACCTTTCAGACCAACCAACCAAAAAGATCTCAAACTCAACCACT	667
Dd		601	AAGCCTACAAAAAAACAACCTTTCAGACCAACCAACCAAAAAGATCTCAAACTCAACCACT	660
OY		668	AAACCAAGGAAGTAGCCACCAACCAACCCACAGAAAGACCAACCATCAACACCAACCAAA	727
Dd		661	AAACCAAGGAAGTAGCCACCAACCAACCCACAGAAAGACCAACCATCAACACCAACCAAA	720
OY		728	ACAAATCATCACTACACTCTCTCAGCAACAACACACAGAGAAATCCAAAATCCAACTCAAGT	787
Dd		721	ACAAATCATCACTACACTCTCTCAGCAACAACACACAGAGAAATCCAAAATCCAACTCAAGT	780
OY		788	CAATGGAATCTTCACTCACTCACTCTCTCCGAAGGCATCTTAAGCCCTTTCAGTCTCC	847
Dd		781	CAATGGAATCTTCACTCACTCACTCTCTCCGAAGGCATCTTAAGCCCTTTCAGTCTCC	840
OY		848	ACAACATCCGACACCACTCAACCAACCTCATCTCAACCCACAACAACAGGCGAC	901
Dd		841	ACAACATCCGACACCACTCAACCAACCTCATCTCAACCCACAACAACAGGCGAC	894
RESULT 10				
US-08-852-344D-28				
Sequence 28, Application US/08852344D				
Patent No. 6017539				
GENERAL INFORMATION:				
APPLICANT: KLEIN, Michel H				
APPLICANT: DU, Run-Pan				
APPLICANT: EWAASYSHN, Mary E				
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION				
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY				
TITLE OF INVENTION: SYNCTIAL VIRUS				
NUMBER OF SEQUENCES: 37				



QY	8	ATGTCCAAAACAAAGACCAAGGACGGCTAAAGCAACTGAAAGAGCCCTGGACACTCTC	67
Db	1	ATGTCCAAAACAAAGACCAAGGACCGGCTAAAGCAACTGAAAGAGCCCTGGACACTCTC	60
QY	68	AATCATTTATTTATTCATATCATCGGGGCTTTATTAAGTTAAATCTTAAATCTGTAGACAA	127
Db	61	AATCATTTATTTATTCATATCATCGGGGCTTTATTAAGTTAAATCTTAAATCTGTAGACAA	120
QY	128	ATCAACATTAATCCATTTCTGGCAATATATATCTCAACTTCACTTAATATTAACGCAATCATTA	187
Db	121	ATCAACATTAATCCATTTCTGGCAATATATATCTCAACTTCACTTAATATTAACGCAATCATTA	180
QY	188	TTTCATAGCTGGGCAACCAACAAGGTACACATTAACAATGCAATATACAGAATGCAACA	247
Db	181	TTTCATAGCTGGGCAACCAACAAGGTACACATTAACAATGCAATATACAGAATGCAACA	240
QY	248	AGCCAGATCAGAGAACACACACCCCAACATACCTCACTAGAGATCTTAGCTTGGAAATGAGC	307
Db	241	AGCCAGATCAGAGAACACACACCCCAACATACCTCACTAGAGATCTTAGCTTGGAAATGAGC	300
QY	308	TTCTCCATCTGTCTGAAATTTACATCATCAACACGACAGCATCTGTGTTTCAACAACGCA	365
Db	301	TTCTCCATCTGTCTGAAATTTACATCATCAACACGACAGCATCTGTGTTTCAACAACGCA	360
QY	368	GGAGTCAAGTCAAACTGTGCAACCCACCAACAGTCAAGCTTAAAGCAACAACAACCCAA	427
Db	361	GGAGTCAAGTCAAACTGTGCAACCCACCAACAGTCAAGCTTAAAGCAACAACAACCCAA	420
QY	428	ACACAACCCACAGACGCCATCTACAAAAACAAGCCAAACAAACGACCAACAAMCCCAAT	487
Db	421	ACACAACCCACAGACGCCATCTACAAAAACAAGCCAAACAAACGACCAACAAMCCCAAT	480

TELEFAX: (416) 595-1163  
TELEX: 065-34567 STYBAC

INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-344-639E-28

Query Match 97.2%; Score 894; DB 3; Length 894;  
Best Local Similarity 100.0%; Pred. No. 6e-236;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 8 ATGTCCAAAACAGAGGACGACGCGGTAGACACTGAGAAAGACCTGGGACACTCTC 67
Db 1 ATGTCCAAAACAGAGGACGACGCGGTAGACACTGAGAAAGACCTGGGACACTCTC 60
QY 68 AATCATTTTATTTATTCATCATCATCGGGCTATATTAAGTTAAATCTTAATCTGTAGACAA 127
Db 61 AATCATTTTATTTATTCATCATCATCGGGCTATATTAAGTTAAATCTTAATCTGTAGACAA 120
QY 128 ATCAATTTATCCATTTCTGGCAATGATATCTCACTTCACTTATATTTACAGCCATCATA 187
Db 121 ATCAATTTATCCATTTCTGGCAATGATATCTCACTTCACTTATATTTACAGCCATCATA 180
QY 188 TTCTATGACCTCGGCAACCAAGTACACACTAACAACTGCAATCATACAGATGCAACA 247
Db 181 TTCTATGACCTCGGCAACCAAGTACACACTAACAACTGCAATCATACAGATGCAACA 240
QY 248 AGCCAGATCAAGAACACACACCCCAACATACCTCAGAGATCTCAGCTTGGATCAGC 307
Db 241 AGCCAGATCAAGAACACACCCCAACATACCTCAGAGATCTCAGCTTGGATCAGC 300
QY 308 TTCTCAATCTGTCTGAAATTTATCATCAACAACCCACCACCTACTAGCTTCAACACACCA 367
Db 301 TTCTCAATCTGTCTGAAATTTATCATCAACAACCCACCACCTACTAGCTTCAACACACCA 360
QY 368 GGAGTCAAGTCAACCTGCAACCCACACAGTCAAGACTTAAACCAACACACACCCAA 427
Db 361 GGAGTCAAGTCAACCTGCAACCCACACAGTCAAGACTTAAACCAACACACACCCAA 420
QY 428 ACACAAACCCGCAAGGCCACTACAAAACAGCCAAACCAACACCAACCAACCAACCAAT 487
Db 421 ACACAAACCCGCAAGGCCACTACAAAACAGCCAAACCAACACCAACCAACCAACCAAT 480
QY 488 AATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTTGACGATATGACGACCAACATCA 547
Db 481 AATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTTGACGATATGACGACCAACATCA 540
QY 548 ACCTGCTGGCTATCTGCAAAAGATTCACAAACAAAACACAGAAAGAAACCCACACG 607
Db 541 ACCTGCTGGCTATCTGCAAAAGATTCACAAACAAAACACAGAAAGAAAGAAACCCACACG 600
QY 608 AAGGCTTCAAAAAAACCAACCTTCAAGACAAACAAAACAAATCTCAAACTCAAAACACT 667
Db 601 AAGGCTTCAAAAAAACCAACCTTCAAGACAAACAAAACAAATCTCAAACTCAAAACACT 660
QY 668 AAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACATCAACACACACCAA 727
Db 661 AAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACATCAACACACACCAA 720
QY 728 ACAACATCAGCACTACAGCTGCTCAGCAACAAACCAAGAAATCCAAAACCTCACAAGT 787
Db 721 ACAACATCAGCACTACAGCTGCTCAGCAACAAACCAAGAAATCCAAAACCTCACAAGT 780
QY 788 CAATGGAACCTTCCACTACCTCCTCGAAGGCAATTAAGCCCTTCTCAAGTCTCC 847
Db 781 CAATGGAACCTTCCACTACCTCCTCGAAGGCAATTAAGCCCTTCTCAAGTCTCC 840
QY 848 ACAACATCAGCAACCCATCAACAACCTCTCTCCACCAACCAACACAGCCGAG 901
Db 841 ACAACATCAGCAACCCATCAACAACCTCTCTCCACCAACCAACACAGCCGAG 894
```

RESULT 12  
US-08-801-898A-23  
Sequence 23, Application US/08801898A  
Patent No. 5998602

GENERAL INFORMATION:  
APPLICANT: Torrence, Paul F.  
APPLICANT: Silverman, Robert H.  
APPLICANT: Cirino, Nick M.  
APPLICANT: Li, Guyling  
APPLICANT: Xiao, Wei  
TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE  
TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/801,898A  
FILING DATE: 18-FEB-1997  
CLASSIFICATION: 514

## ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.  
REGISTRATION NUMBER: 28,462  
REFERENCE/DOCKET NUMBER: 8656-009  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741

## INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:  
LENGTH: 15222 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

## FEATURE:

NAME/KEY: RSV-A2  
LOCATION: 1...15222  
OTHER INFORMATION:

US-08-801-898A-23

Query Match 92.8%; Score 853.4; DB 2; Length 15222;  
Best Local Similarity 95.5%; Pred. No. 2.9e-224;  
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 1 TCGAAGATGTCGCAAAACAGGACCAACGACGCTAAGACACTAGAAAAGACCTGGGA 60
Db 4681 TCGAAGATGTCGCAAAACAGGACCAACGACGCTAAGACACTAGAAAAGACCTGGGA 4740
QY 61 CACTCTCAATCATTTATTTATTCATATCATTCGGGCTTATTAAGTTAAATCTTAATCTGT 120
Db 4741 CACTCTCAATCATTTATTTATTCATATCATTCGGGCTTATTAAGTTAAATCTTAATCTGT 4800
QY 121 AGCACAATACATTAATTCATTTCTGGCAATGATATCTCAACTCACTTATTAATTAACAGC 180
Db 4801 AGCACAATACATTAATTCATTTCTGGCAATGATATCTCAACTCACTTATTAATTAACAGC 4860
QY 181 CATCATATTCATAGGCTCGGCAACCAACAAAGTCACTAACAATGCAATCATACAGA 240
Db 4861 CATCATATTCATAGGCTCGGCAACCAACAAAGTCACTAACAATGCAATCATACAGA 4920
QY 241 TGCACAGCCGATCAAGAACACACCCCAACATACCTCACTAGATCTCAGCTTGG 300
Db 4921 TGCACAGCCGATCAAGAACACACCCCAACATACCTCACTAGATCTCAGCTTGG 4980
```

QY 301 AATGAGCTTCTCCAAATCTGTCTGAAATTTACATCACAAAACCCACCATCTAGTTCAAC 360  
 Db 4981 AATGAGCTTCTCTATCTGATCGTGAATTTACATCAACAAATACACCCATCTAGTTCAAC 5040  
 QY 361 AAGACGAGATTCAGATTAACCTGCAACCCACAGAGTAAAGCTTAAACCAACAC 420  
 Db 5041 AACACGAGATTCAGATTAACCTGCAACCCACAGAGTAAAGCTTAAACCAACAC 5100  
 QY 421 AACCAACACACACCCAGCAAGCCCTACACAAAACAGCCAAACACCAACCAACAA 480  
 Db 5101 AACCAACACACACCCAGCAAGCCCTACACAAAACAGCCAAACACCAACCAACAA 5160  
 QY 481 ACCCAATATGATTTTCACTTGAAGTGTCTTAACTTTGACCTGCGATATGACGAA 540  
 Db 5161 ACCCAATATGATTTTCACTTGAAGTGTCTTAACTTTGACCTGCGATATGACGAA 5220  
 QY 541 CAATCCAAACCTGTGGGCTATCTGCAAAACATATCCAAACAAAACCAAGAAAGAAAC 600  
 Db 5221 CAATCCAAACCTGTGGGCTATCTGCAAAACATATCCAAACAAAACCAAGAAAGAAAC 5280  
 QY 601 CACCACCAAGCTTCAAAAACCAACCTTCAAGACACCAAGCCACAGAAAGACCAATCAACCTCA 660  
 Db 5281 CACCACCAAGCTTCAAAAACCAACCTTCAAGACACCAAGCCACAGAAAGACCAATCAACCTCA 5340  
 QY 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCACAGAAAGACCAATCAACCTCA 720  
 Db 5341 AACCACTAAACCAAGAGTACCCACCAACCAAGCCACAGAAAGACCAATCAACCTCA 5400  
 QY 721 CACCAAAACCAATCACAATCACTAGCTGTCACCAACCAACCAAGAAATTCAAACT 780  
 Db 5401 CACCAAAACCAATCACAATCACTAGCTGTCACCAACCAACCAAGAAATTCAAACT 5460  
 QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840  
 Db 5461 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 5520  
 QY 841 AGTCTCACAACATCCGAGCACCACATACAAACCTTATCTCAACCAACCAACAGGCA 900  
 Db 5521 AGTCTCACAACATCCGAGCACCACATACAAACCTTATCTCAACCAACCAACAGGCA 5580  
 QY 901 GTAGTTATTAAAAAAA 919  
 Db 5581 GTAGTTATTAAAAACATA 5599

RESULT 13  
 US-08-962-690-12/c  
 ; Sequence 12, Application US/08962690  
 ; Patent No. 6214805  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Torrence, Paul F.  
 ; APPLICANT: Silverman, Robert H.  
 ; APPLICANT: Clirino, Nick M.  
 ; APPLICANT: Li, Guiying  
 ; APPLICANT: Xiao, Wei  
 ; APPLICANT: Player, Mark R.  
 ; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE OLIGONUCLEOTIDES  
 ; TITLE OF INVENTION: EFFECTIVE TO TREAT RSV INFECTIONS  
 ; FILE REFERENCE: 8656-019  
 ; CURRENT APPLICATION NUMBER: US/08/962,690  
 ; EARLIER APPLICATION NUMBER: 08/801,896  
 ; EARLIER FILING DATE: 1997-02-14  
 ; EARLIER APPLICATION NUMBER: 60/011,725  
 ; NUMBER OF SEQ ID NOS: 40  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 12  
 ; LENGTH: 15222  
 ; TYPE: DNA  
 ; ORGANISM: respiratory syncytial virus  
 US-08-962-690-12

Query Match 92.88; Score 853.4; DB 4; Length 15222;  
 Best Local Similarity 95.58; Pred. No. 2,9e-224;  
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
 QY 1 TGCACATGTCGCAAAAAGAGCAAGCACCCTGTAAGACACTGAAAAGACCTGGGA 60  
 Db 10542 TGCACATGTCGCAAAAAGAGCAAGCACCCTGTAAGACACTGAAAAGACCTGGGA 10483  
 QY 61 CACTCTCAATCATTTATTTATTCATATCATATGCGCTTATATTAATTTAAATCTGT 120  
 Db 10482 CACTCTCAATCATTTATTTATTCATATCATATGCGCTTATATTAATTTAAATCTGT 10423  
 QY 121 AGCAAAATCATCATTTATTCATTTGCGCAATGATATTCGCAACTTCACTTA"AAATACAGC 180  
 Db 10422 AGCAAAATCATCATTTATTCATTTGCGCAATGATATTCGCAACTTCACTTA"AAATACAGC 10363  
 QY 181 CATCATATTCATGCTCTGGCAAAACCAAGTACACATACAGTACAGTACATACAGTA 240  
 Db 10362 CATCATATTCATGCTCTGGCAAAACCAAGTACACATACAGTACAGTACATACAGTA 10303  
 QY 241 TGCAACAAAGCCAGATCAAGAAACAACCCCAATACCTCTCAGATCTCAAGCTTGG 300  
 Db 10302 TGCAACAAAGCCAGATCAAGAAACAACCCCAATACCTCTCAGATCTCAAGCTTGG 10243  
 QY 301 AATGAGCTTCTCCAAATCTGTCTGAAATTTACATACAAACCAACCAACCTACTAGCTTCAAC 360  
 Db 10242 AATGAGCTTCTCCAAATCTGTCTGAAATTTACATACAAACCAACCAACCTACTAGCTTCAAC 10183  
 QY 361 AACACAGAGTCAAGTCAACCAACCTGCAACCAACAGTCAAGTCAAGTCAAGTCAAG 420  
 Db 10182 AACACAGAGTCAAGTCAACCAACCTGCAACCAACAGTCAAGTCAAGTCAAGTCAAG 10123  
 QY 421 AACCAAAACCAACCCAGCAAGCCCACTACAAAACCAACCCCAACCAACCAACCAAC 480  
 Db 10122 AACCAAAACCAACCCAGCAAGCCCACTACAAAACCAACCCCAACCAACCAACCAAC 10063  
 QY 481 ACCCAATATGATTTTCACTTGAAGTGTCTTAACTTTGACCTGCGATATGACGAA 540  
 Db 10062 ACCCAATATGATTTTCACTTGAAGTGTCTTAACTTTGACCTGCGATATGACGAA 10003  
 QY 541 CAATCCAAACCTGTGGGCTATCTGCAAAACATATCCAAACAAAACCAAGAAAGAAAC 600  
 Db 10002 CAATCCAAACCTGTGGGCTATCTGCAAAACATATCCAAACAAAACCAAGAAAGAAAC 9943  
 QY 601 CACCACCAAGCTTCAAAAACCAACCTTCAAGACACCAACCAACCAACCAACCAACCTCA 660  
 Db 9942 CACCACCAAGCTTCAAAAACCAACCTTCAAGACACCAACCAACCAACCAACCAACCTCA 9883  
 QY 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCACAGAAAGACCAATCAACCTCA 720  
 Db 9882 AACCACTAAACCAAGAGTACCCACCAACCAAGCCACAGAAAGACCAATCAACCTCA 9823  
 QY 721 CACCAAAACCAACATCATACATACATACATACATACATACATACATACATACAT 780  
 Db 9822 CACCAAAACCAACATCATACATACATACATACATACATACATACATACATACAT 9763  
 QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTA"GCCCTTCTCA 840  
 Db 9762 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTA"GCCCTTCTCA 9703  
 QY 841 AGTCTCACAACATCCGAGCACCACATACAAACCTTATCTCAACCAACCAACAGGCA 900  
 Db 9702 AGTCTCACAACATCCGAGTACCCATACAAACCTTATCTCAACCAACCAACAGGCA 9643  
 QY 901 GTAGTTATTAAAAAAA 919  
 Db 9642 GTAGTTATTAAAAACATA 9624

RESULT 14  
 US-08-892-403A-1  
 ; Sequence 1, Application US/08892403A

Patent No. 5993824  
GENERAL INFORMATION:  
APPLICANT: Murphy, Brian R.  
APPLICANT: Collins, Peter L.  
APPLICANT: Whitehead, Stephen S.  
APPLICANT: Bukreyev, Alexander A.  
APPLICANT: Juhász, Katalin  
TITLE OF INVENTION: PRODUCTION OF ATTENUATED RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,403A  
FILING DATE: 15-JUL-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,634  
FILING DATE: 23-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,141  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 60/021,773  
FILING DATE: 15-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 17634-000510  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15223 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-892-403A-1

Query Match 92.8%; Score 853.4; DB 2; Length 15223;  
Best Local Similarity 95.5%; Pred. No. 2.9e-224;  
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCACATGTCACAAACAAAGACACACGACCGCTTAAGACACTGAGAAAGACCTGGGA 60  
DB 4682 TGCACATGTCACAAACAAAGACACACGACCGCTTAAGACACTGAGAAAGACCTGGGA 4741  
QY 61 CACCTCAATCATTTATTTATTCATATCATGCGGCTTATATAGTTAAATTTAAATCTGT 120  
DB 4742 CACCTCAATCATTTATTTATTCATATCATGCTTATATAGTTAAATTTAAATCTGT 4801  
QY 121 AGCACAATGACATTTATTCATTTGTCATGATTAATCTCACTTAAATTAATGACG 180  
DB 4802 AGCACAATGACATTTATTCATTTGTCATGATTAATCTCACTTAAATTAATGACG 4861  
QY 181 CACTATATTCATAGCTGCGCAACACCAAGTCACTAACAAGTCAATCATATCAAGA 240  
DB 4862 CACTATATTCATAGCTGCGCAACACCAAGTCACTAACAAGTCAATCATATCAAGA 4921  
QY 241 TGCACACGACGATCAAGAACACCAACCCCAATACCTGACGAGATCCTGAGCTGG 300

DB 4922 TGCACACGACGATCAAGAACACCAACCCCAATACCTGACGAGATCCTGAGCTGG 4981  
QY 301 AATGAGCTTCCATCTCTGTCTGAAATTAATACATACAAACACGACCATAGCTTCAAC 360  
DB 4982 AATGAGCTTCCATCTCTGTCTGAAATTAATACATACAAACACGACCATAGCTTCAAC 5041  
QY 361 AACACGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACACAC 420  
DB 5042 AACACGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACACAC 5101  
QY 421 AACCAACACACACCCGACGAGCCGCTACAAACCAACGCAACCAACCAACCAAC 480  
DB 5102 AACCAACACACACCCGACGAGCCGCTACAAACCAACGCAACCAACCAACCAAC 5161  
QY 481 AACCAATATGATTTTCTCTGCAAGTGTAACTTTGTAACCTGAGCATATGACGCA 540  
DB 5162 AACCAATATGATTTTCTCTGCAAGTGTAACTTTGTAACCTGAGCATATGACGCA 5221  
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAAC 600  
DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAAC 5281  
QY 601 CACCAACGACCTGACCAACCAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660  
DB 5282 CACTACCAAGCCCAACCAACCAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 5341  
QY 661 AACCACTAAACCAAGAGTACCCACACACCAACCAACCAACCAACCAACCAAC 720  
DB 5342 AACCACTAAATCAAGAGTACCCACACCAACCAACCAACCAACCAACCAACCAAC 5401  
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
DB 5402 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5461  
QY 781 CACAAGTCAATGGAACCTTCACTCACTCACTCTCTCGAAGCAATTAAGCTTTCTCA 840  
DB 5462 CACAAGTCAATGGAACCTTCACTCACTCACTCTCTCGAAGCAATTAAGCTTTCTCA 5521  
QY 841 AGCTTCACACATCTCCGACGACCCATCCATCCATCCATCCATCCATCCATCCATCC 900  
DB 5522 AGCTTCACACATCTCCGACGACCCATCCATCCATCCATCCATCCATCCATCCATCC 5581  
QY 901 GTAAGTATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 919  
DB 5582 GTAAGTATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5600

RESULT 15  
US-08-720-132-1  
Sequence 1, Application US/08720132  
Patent No. 6264957  
GENERAL INFORMATION:  
APPLICANT: Collins, Peter L.  
TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,132  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/007,083  
FILING DATE: 27-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 15280-250-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15223 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other  
US-08-720-132-1

Query Match 92.8%; Score 853.4; DB 4; Length 15223;  
Best Local Similarity 95.5%; Pred. No. 2.9e-224;  
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCACATGTCGCAAAACAGAGACGACGCTTAAGACCTAGAAAAGACCTGGGA 60  
DB 4682 TGCACATGTCGCAAAACAGAGACGACGCTTAAGACCTAGAAAAGACCTGGGA 4741  
QY 61 CACTCTCAATCATTTATTTATTCATATCGGGCTTATATAGTTAATCTTAATCTGT 120  
DB 4742 CACTCTCAATCATTTATTTATTCATATCGGGCTTATATAGTTAATCTTAATCTGT 4801  
QY 121 AGCACAATTCATTTATTCATTTCTGGCAATGATTAATCTCACTTCACTTAATTACAGC 180  
DB 4802 AGCACAATTCATTTATTCATTTCTGGCAATGATTAATCTCACTTCACTTAATTACAGC 4861  
QY 181 CATCATATTCATAGCCTGGGCAACCAAGTCACTAACAAGTCACTAATACAGCA 240  
DB 4862 CATCATATTCATAGCCTGGGCAACCAAGTCACTAACAAGTCACTAATACAGCA 4921  
QY 241 TGCACAGCCGAGATCAGACACACCAACCATACCTCACTAGATCTCAGCTTGG 300  
DB 4922 TGCACAGCCGAGATCAGACACACCAACCATACCTCACTAGATCTCAGCTTGG 4981  
QY 301 AATCAGCTTCTCAATCTCTCTGAAATTCATACACAACCCACCATTAAGCTTCAAC 360  
DB 4982 AATCAGCTTCTCAATCTCTCTGAAATTCATACACAACCCACCATTAAGCTTCAAC 5041  
QY 361 AACACGAGGAGTCAAGTCAAACTGCAACCCACAGTCAAGCTTAATAACACACAAC 420  
DB 5042 AACACGAGGAGTCAAGTCAAACTGCAACCCACAGTCAAGCTTAATAACACACAAC 5101  
QY 421 AACCCAAACACACCCAGCAGCAGCTACAAACAGCCCAAAACCAAAACCAACCA 480  
DB 5102 AACCTCAAAACACCCAGCAGCAGCTACAAACAGCCCAAAACCAAAACCAACCA 5161  
QY 481 ACCCAATATGATTTTCACTTGAAGTGTATTACTTTGACCTTCAGCATATGCAACA 540  
DB 5162 ACCCAATATGATTTTCACTTGAAGTGTATTACTTTGACCTTCAGCATATGCAACA 5221  
QY 541 CAATCCAACTTGTGGCTATCTGCAAAAAGATACCAAAACCAAAACCAAGAAACAA 600  
DB 5222 CAATCCAACTTGTGGCTATCTGCAAAAAGATACCAAAACCAAAACCAAGAAACAA 5281  
QY 601 CACCAACCAAGCTACAAAAAACCAACTTCAGACAAACCAAAAGATCTCAAACTCA 660  
DB 5282 CACTACCAAGCTACAAAAAACCAACTTCAGACAAACCAAAAGATCTCAAACTCA 5341  
QY 661 AACCACTAAACCAAGGAGTACCCACACCAAGGCCACAGAAGGCCAACCATCAACAC 720  
DB 5342 AACCACTAAATCAAGGAGTACCCACACCAAGGCCACAGAAGGCCAACCATCAACAC 5401  
QY 721 CACCAAAACCAACATCACACTACACTGCTCACCAACCAACACACAGGAAATCCAAAAC 780  
DB 5402 CACCAAAACCAACATCACACTACACTGCTCACCAACCAACACAGGAAATCCAAAAC 5461

QY 781 CACAAGTCAATGAAACCTTCCACTCAACCTCTCCGAAAGGCAATCTTAAGCTTTCTCA 840  
DB 5462 CACAAGTCAATGAAACCTTCCACTCAACCTCTCCGAAAGGCAATCTTAAGCTTTCTCA 5521  
QY 841 AGCTTCCACACATCCGAGCACCACATCAACCTTCATCTCCACCAACACACAGGCA 900  
DB 5522 AGCTTCCACACATCCGAGCACCACATCAACCTTCATCTCCACCAACACACAGGCA 5581  
QY 901 GTAGTTATTAATAAAAAA 919  
DB 5582 GTAGTTACTTAATAAACATA 5600

Search completed: May 1, 2003, 12:35:08  
Job time : 63.8257 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 11:59:58 ; Search time 94.5321 Seconds  
(without alignments)  
11499.705 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacaatgcacaaaca.....gtagtattataaaaaaaa 920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PC1\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PC105\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	853.4	92.8	15223 9	US-09-847-173-1
2	637.4	69.3	696 10	US-09-844-645-4
3	85.8	9.3	793 10	US-09-878-574-4304
4	78.2	8.5	1635 10	US-09-864-761-20241
5	78.2	8.5	1973 10	US-09-864-761-3471
6	76.6	8.3	576 10	US-09-878-574-4296
7	76.4	8.3	529 10	US-09-883-965-2109
8	76.4	8.3	1030 10	US-09-878-574-4306
9	76.4	8.3	1075 10	US-09-864-761-19241
10	76.4	8.3	1403 10	US-09-864-761-2513
11	75	8.2	577 10	US-09-878-574-4312
12	74	8.0	639 10	US-09-878-574-4316
13	70	7.6	576 10	US-09-864-761-26582
14	69	7.5	1428 9	US-10-001-857-79
15	68.6	7.5	1423 9	US-10-001-857-100
16	68.2	7.4	491 10	US-09-878-574-4301
17	67.6	7.3	584 10	US-09-878-574-4310
18	63.6	6.9	555 10	US-09-864-761-9605
19	63.4	6.9	766 10	US-09-878-574-4344

c 20	63	6.8	528	10	US-09-878-574-4305	Sequence 4305, Ap
c 21	63	6.8	1267	12	US-10-001-843-45	Sequence 45, Appl
c 22	61.2	6.7	446	10	US-09-960-352-3400	Sequence 3400, Ap
c 23	61	6.6	448	10	US-09-878-574-4308	Sequence 4308, Ap
c 24	60.2	6.5	341	10	US-09-960-352-12302	Sequence 12302, A
c 25	59	6.4	545	10	US-09-878-574-4299	Sequence 4299, Ap
c 26	57.6	6.3	1557	9	US-10-001-873-1	Sequence 1, Appl
c 27	57.4	6.2	425	10	US-09-960-352-4010	Sequence 4010, Ap
c 28	57.4	6.2	1938	10	US-09-834-975-1039	Sequence 1039, Ap
c 29	57.2	6.2	119596	9	US-10-270-336-3	Sequence 3, Appl
c 30	57	6.2	446	10	US-09-864-761-20699	Sequence 20699, A
c 31	56	6.1	155074	9	US-10-026-188-6	Sequence 6, Appl
c 32	55.6	6.0	283	10	US-09-960-352-9095	Sequence 9095, Ap
c 33	54	5.9	415	9	US-10-025-380-1058	Sequence 2223, Ap
c 34	53	5.8	15720	10	US-09-922-217-1058	Sequence 1058, Ap
c 35	53	5.8	15720	10	US-09-833-263-1058	Sequence 1058, Ap
c 36	53	5.8	15720	10	US-09-864-761-768	Sequence 768, Ap
c 37	51.8	5.6	479	10	US-09-864-761-19531	Sequence 19531, A
c 38	51.8	5.6	830	10	US-09-825-294-53	Sequence 53, Appl
c 39	51.4	5.6	396	9	US-10-083-357-341	Sequence 341, Ap
c 40	51.4	5.6	462	9	US-09-960-352-7186	Sequence 7186, Ap
c 41	51.4	5.6	463	10	US-09-864-761-6690	Sequence 6690, Ap
c 42	51.2	5.5	600	10	US-09-960-352-1009	Sequence 1009, Ap
c 43	51	5.5	439	10	US-09-960-352-5785	Sequence 5785, Ap
c 44	50.8	5.5	516	10		
c 45	50.8	5.5	516	10		

#### ALIGNMENTS

RESULT 1  
US-09-847-173-1  
Sequence 1, Application US/09847173  
Publication No. US20020182228A1  
GENERAL INFORMATION:  
APPLICANT: Collins, Peter L.  
TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/847,173  
FILING DATE: 03-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/720,132  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 15280-250-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15223 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-847-173-1

Query Match 92.8%; Score 853.4; DB 9; Length 15223;  
Best Local Similarity 95.5%; Pred. No. 5,2e-211;  
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 1 TCGAATCATGTCCAAAAGAGGACGACCGCTAGACATAGAAAAGACCTGGGA 60
    |||||||
DB 4682 TCGAATCATGTCCAAAAGAGGACGACCGCTAGACATAGAAAAGACCTGGGA 4741
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAATCTGT 120
    |||||||
DB 4742 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAATCTGT 4801
QY 121 AGCAAAATTCATTTATTCATTTCTGGCAATGATATTCACCTCATTAATTAACAGC 180
    |||||||
DB 4802 AGCAAAATTCATTTATTCATTTCTGGCAATGATATTCACCTCATTAATTAATTCAGC 4861
QY 181 CATCATATTCATATGCTGGGCAAGCAAGTCACTATGACACTGCAATCATATCAAGA 240
    |||||||
DB 4862 CATCATATTCATATGCTGGGCAAGCAAGTCACTATGACACTGCAATCATATCAAGA 4921
QY 241 TGCACAGGACGATCAAGACACACACCCCAATACCTACTCTAGATCTCTAGCTTG 300
    |||||||
DB 4922 TGCACAGGACGATCAAGACACACACCCCAATACCTACTCTAGATCTCTAGCTTG 4981
QY 301 AATCAGCTTCTGCAATCTGCTGTAATTTACATACAAACACACCATCTACTGCTGAC 360
    |||||||
DB 4982 AATCAGCTTCTGCAATCTGCTGTAATTTACATACAAATCCATCTACTGCTGCTGAC 5041
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACACAC 420
    |||||||
DB 5042 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACACAC 5101
QY 421 AACCCAAACACACACACAGGACCTCTACAAACACAGCCCAAAACACACACACAA 480
    |||||||
DB 5102 AACCTCAACACACACACAGGACCCACACACAAACACAGCCCAAAACACACACAA 5161
QY 481 ACCCAATTAATGATTTTCACTTCAAGTGTATTAATCTTTGACCTGACGATATGACGAA 540
    |||||||
DB 5162 ACCCAATTAATGATTTTCACTTCAAGTGTATTAATCTTTGACCTGACGATATGACGAA 5221
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACAAAACCAAGAAAGAAAC 600
    |||||||
DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACAAAACCAAGAAAGAAAC 5281
QY 601 CACCACCAAGCTTACAAAAAACCAACCTTCAAGACACCAAAAAAAGATCTCAAACTCA 660
    |||||||
DB 5282 CACTACCAAGCTTACAAAAAACCAACCTTCAAGACACCAAAAAAAGATCTCAAACTCA 5341
QY 661 AACCACTAAACCAAGAGATACCCACACACAGCCACAGAGAGCCAAACATCAACAC 720
    |||||||
DB 5342 AACCACTAAATCAAGAGAGATACCCACACACAGCCACAGAGAGCCAAACATCAACAC 5401
QY 721 CACCAAAACCAACATCACTACTGCTACCAACACACAGGAGAAATCTCAAAACT 780
    |||||||
DB 5402 CACCAAAACCAACATCACTACTGCTACCAACACACAGGAGAAATCTCAAAACT 5461
QY 781 CACAAGCAATGGAACCTTCAACCTCAACCTGCGAAGGCAATCTAAGCCCTCTCA 840
    |||||||
DB 5462 CACAAGCAATGGAACCTTCAACCTCAACCTGCGAAGGCAATCTAAGCCCTCTCA 5521
QY 841 AGTCTCACAACATCCGAGCAGCCCAATCAACCTCATCTCCACCAACACACAGGCA 900
    |||||||
DB 5522 AGTCTCACAACATCCGAGCAGCCCAATCAACCTCATCTCCACCAACACACAGGCA 5581
QY 901 GTAGTATTAAAAAAA 919
    |||||||
DB 5582 GTAGTATTAAAAACATA 5600
    |||||||
```

RESULT 2

```
US-09-844-645-4
: Sequence 4, Application US/09844645
: Patent No. US20020102242A1
: GENERAL INFORMATION:
:   APPLICANT: Billies, David E.
:             McDaniel, Larry S.
:             Curriel, David T.
: TITLE OF INVENTION: COMPOSITION AND METHODS FOR
:                   ADMINISTERING PNEUMOCOCCAL DNA
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
:   ADDRESS: Curriel, Morris & Safford, P.C.
:   STREET: 530 Fifth Avenue
:   CITY: New York
:   STATE: New York
:   COUNTRY: U.S.A.
:   ZIP: 10036
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/844,645
:   FILING DATE: 27-Apr-2001
:   CLASSIFICATION: <unknown>
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: US 08/759,505
:     FILING DATE: 04-DEC-1996
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Frommer, William S.
:     REGISTRATION NUMBER: 25,506
:     REFERENCE/DOCKET NUMBER: 454312-2450
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (212) 840-3333
:   INFORMATION FOR SEQ ID NO: 4:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 696 base pairs
:       TYPE: nucleic acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:       MOLECULE TYPE: DNA (genomic)
:       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
:
: US-09-844-645-4
:
: Query Match 69.3%; Score 637.4; DB 10; Length 696;
: Best Local Similarity 96.1%; Pred. No. 1.2e-155;
: Matches 664; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
:
: QY 8 ATGTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 67
:   |||||||
: DB 1 ATGTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 60
: QY 68 AATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGCAAA 127
:   |||||||
: DB 61 AATCATTTATTATTCATTCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGCAAA 120
: QY 128 ATACATTATTCATTCATTCGCAATGATATTCOAATCTTACTTAATTAATACAGC-ATCATA 187
:   |||||||
: DB 121 ATACATTATTCATTCATTCGCAATGATATTCOAATCTTACTTAATTAATTAATGACATATA 180
: QY 188 TTGATAGCTTGGCAAAACCAAGTCACTAAGTCAATGCAATCATTAAGATGACAA 247
:   |||||||
: DB 181 TTGATAGCTTGGCAAAACCAAGTCACTAAGTCAATGCAATCATTAAGATGACAA 240
: QY 248 AGCCAGATCAAGACACACACCCCAATACCTCACTGATCTCTGAGCTTGAGATGAGC 307
:   |||||||
: DB 241 AGCCAGATCAAGACACACACCCCAATACCTCACTGATCTCTGAGCTTGAGATGAGT 300
: QY 308 TTCTCCAAATCTGTCTGAAATTACATCAAAAACCAACCATATGATGCTTCAACAACACA 367
:   |||||||
: DB 301 CCCTCTAATCGCTGGAATTATGATCAAAAATCAACCATATGATGCTTCAACAACACA 360
```







```

RESULT 6
US-09-878-574-4296/C
; Sequence 4296 Application US/09878574
; Patent No. US20020110546A1
GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535

```

```

: RESULT 7
: US-09-983-965-2109/c
: Sequence 2109, Application US/09983965
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nenping
: APPLICANT: Byatt, John C.
: APPLICANT: Mathalagan, Nagappan
: TITLE OF INVENTION: NOCLETIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: FILE REFERENCE: 37-21(10297)C
: CURRENT APPLICATION NUMBER: US/09/983,965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US 09/465,231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/113,678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 2109
: LENGTH: 529
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (335)
: OTHER INFORMATION:

```

OTHER INFORMATION: Clone ID: 34-LIB3057-015-Q1-K1-A6  
US-09-983-965-2109

Query Match 8.3%; Score 76.4; DB 10; Length 529;  
Best Local Similarity 47.9%; Pred. No. 3.8e-10;  
Matches 218; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

QY 324 AATTATCATCACAACCCACCATCTAGCTTCAACAACCCAGAGTCAAGTCAAC 383  
DB 504 AA 445  
QY 384 TCAACCCACAGTCAAGACTTAAACACACACACACACACACACACACAC 443  
DB 444 AA 385  
QY 444 CCAGTCAACACAGCCCAACACACACACACACACACATATGTTTCCTG 503  
DB 384 AA 325  
QY 504 AAGTGTCTTACTTTGACCTGAGCATATGACACACATCTCACTGCTGCTATCT 563  
DB 324 AA 265  
QY 564 GCAGAAATATACCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 623  
DB 264 AA 205  
QY 624 CACCTTCAAGCAACCAACCAACCTCAACCTCAACCTCAACCAAGAGAGTAC 683  
DB 204 AA 145  
QY 684 CCACCAACAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 743  
DB 144 AA 85  
QY 744 CACTGCTACCAACACACACAGAGAAATCCAAA 778  
DB 84 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50

## RESULT 8

US-09-878-574-4306/c  
Sequence 4306, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 4306  
LENGTH: 1030  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(1030)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-D11  
US-09-878-574-4306

Query Match 8.3%; Score 76.4; DB 10; Length 1030;  
Best Local Similarity 41.0%; Pred. No. 5.3e-10;  
Matches 285; Conservative 0; Mismatches 399; Indels 11; Gaps 2;  
QY 202 AACCACAAAGTCACTAATGCAATCAATCAAGAGCAACAAAGCAGATCAAGAA 261

DB 793 ACAACACAAC 734  
QY 262 CACAAACCCCAACATACCTCAGATCTCAGATCTGGAATGAGTCTCCATCTGTC 321  
DB 733 AACCAAC 674  
QY 322 TGAATTTACATCAACCAACACACATCTAGCTTCAACCAACACAGAGTCAAGTCA 381  
DB 673 ANA-----MCAACCAACACACACACACACACACACACACACACACACAC 622  
QY 382 CCTGCAACCCCAACACATCAAGACTTAAACACACACACACACACACACACAC 441  
DB 621 ACAACCAACCAACACACACACACACACACACACACACACACACACACAC 562  
QY 442 GCCCATACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 501  
DB 561 NCACANNCNANNCACACACACACACACACACACACACACACACACACACAC 502  
QY 502 CGAAGTGTTAAGTTTGTACCTGACGATATGACGACACATCTCAACTGCTGGGCTAT 561  
DB 501 ACCCCCCCAGAC 442  
QY 562 CTGCAAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 621  
DB 441 CAC 382  
QY 622 ACCAAGCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 681  
DB 381 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 325  
QY 682 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 741  
DB 324 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 265  
QY 742 TACACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 801  
DB 264 NACACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNAC 205  
QY 802 CCACTCAACCTCTCTCGAAGCAATCTTAAGCCCTTCTCAAGTCTCCACACATCCGAGCA 861  
DB 204 CNANACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 145  
QY 862 CCATCAACCAACCTCTCAATCTCCACCAACCAACCAACCAACCAACCA 896  
DB 144 NNANCNANNN 110

## RESULT 9

US-09-864-761-19241/c  
Sequence 19241, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aesomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30

```

? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 19241
? LENGTH: 1075
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AL078472.1
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
? OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 43
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
? OTHER INFORMATION: EST_HUMAN HIT: AV739739.1, EVALU 1.00e+00
? OTHER INFORMATION: NT HIT: AL163201.2, EVALU 2.00e-19
US-09-864-761-19241

Query Match      8.3%: Score 76.4; DB 10; Length 1075;
Best Local Similarity 44.5%: Pred. No. 5,5e-10;
Matches 343; Conservative 0; Mismatches 426; Indels 1; Gaps 1;

```

```

QY 488 AATGATTTTCACTTGGAGTGTTTACTTTGTACCTTCGACATATGACGCGAATCA 547
DB 549 ACTTACCACCACCACCCACCCACCATATCACTAAATTAACACGACCCACCCACCC 490
QY 548 ACCCTGCTGGCTATCTGCAAAAGAAATACCAAAACAAAGAAAGAAAGCAACCC 607
DB 489 ACCATCATCACCATCACCATCACTACGACGACACACACACCATCATTACCACTAC 430
QY 608 AAGCTTCAAAAAAACAACCTTCAAGACCAACCAAAAAAGATCTCAAAACCTCAACCA 667
DB 429 ACCACCTACACCATCACCACGACGACGACGACGACGACGACGACGACGACGAC 370
QY 668 AAACCAAGGAAGTATCCGACCCACCAAGCCCAAGAGGCAAGCCATCAACACCA 727
DB 369 ACCACTTACATCACCACCATCACCACCAACCATCACTACCATCACCACCATCACC 310
QY 728 ACAACATCAACAACCTGCTGACCAACCAACGACGACGACGACGACGACGACGAC 787
DB 309 ACCACATCACCACCATCACCATCACCACGACGACGACGACGACGACGACGACGAC 250
QY 788 CAATGGAACCTTCACTCAACCTCTCGGAGGCA-ATCTAAGCCCTTCTCAAGTCTC 846
DB 249 ACCACATCACCACCATCACCACCATCACCACGACGACGACGACGACGACGACGAC 190
QY 847 CACAACATCGGAGCACCACCATCAACACCTGATCTGCAACCCCAACCAACAC 896
DB 189 ACCACCATCACCACCATCACCACCATCACCACGACGACGACGACGACGACGAC 140

RESULT 10
US-09-864-761-2513/c
? Sequence 2513, Application US/09864761
? Patent No. US20020048763A1
? GENERAL INFORMATION:
? APPLICANT: Penn, Sharon G.
? APPLICANT: Rank, David R.
? APPLICANT: Hanzel, David K.
? APPLICANT: Chen, Wensheng
? TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
? FILE REFERENCE: Aeomica-X-1
? CURRENT APPLICATION NUMBER: US/09/864,761
? PRIOR FILING DATE: 2001-05-23
? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: GB 24263.6
? PRIOR FILING DATE: 2000-10-04
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670

```



```
RESULT 12
US-09-878-574-4316/c
; Sequence 4316; Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIORITY FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4316
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(639)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-H11
US-09-878-574-4316

Query Match
Best Local Similarity 49.18; Pred. No. 1.8e-09;
Matches 223; Conservative 0; Mismatches 230; Indels 1; Gaps 1;

QY 324 AAATTACATCAACAACCCATCTAGCTTACAGACACCGAGTCAAGTCAACC 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 AAACAACACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 TGCAACCCACACAGTCAAGCTAAAAACACACACACACACACACACAC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 ACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 444 CCATACAAAACAGCCAAACAAACAAACAAACAAACAAACAAACAAAC 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 ACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 AAGTGTTCACCTTGACCTGCAGCATATGCAGCAACATCCAACTGCTG 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 AACACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 GCAAAAGCATACCAAAACAAAACAGGAAAGAAAACCCACCAAGCTTCA 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 CAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 624 CAACCTTCACACAAACAAACAAACAAACAAACAAACAAACAAACAA 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 ACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 CCACCACCAAGCCCAACAGAGCCCAACCATCAACACCCCAACCAACAT 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 ACACCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 744 CACTGCTTACCAACACACACAGGAAATCCAA 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 CAACCAACAAACAAACCAACCAACCAACCAACCAACCAACCAAC 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-864-761-26582/c
; Sequence 26582; Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26582
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000111.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P44836, EVALUUE 1.00e-12
US-09-864-761-26582

Query Match
Best Local Similarity 7.68; Score 70; DB 10; Length 576;
Matches 170; Conservative 0; Mismatches 125; Indels 5; Gaps 2;

QY 427 AACACAAACGAGAGCCCTACAAACAAACGACCCCAACCAACCAACCA 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 AAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 TAAATGATTTCACTGCAAGGTGTTAACTTGTACCTGCAGCATATGCA 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 CAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 AACCTGTGGGTATCTGCAAAAGAAATACCAACCAACCAACCAACCA 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 CAAAGCTACAAACCAACCTTCAAGACACCAACCAACCAACCAACCA 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





Db 1350 AAAACATTAACACATATAAAAAATTAATAAAAAATAAACACCAAAAAATA 1409  
Oy 803 CAC 805  
Db 1410 AAC 1412

Search completed: May 1, 2003, 13:43:09  
Job time : 108.532 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:08:32 ; Search time 1137.76 Seconds  
(without alignments)  
13095.757 Million cell updates/sec

Title: US-09-462-816-1  
Perfect score: 920  
Sequence: 1 tgcgaacatgtccaaaaa.....gtagtattataaaaaaaa 920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estdb:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	13.0	895	17	CNS0071A
2	104.6	11.4	946	17	AG127412 Pan trogl
3	103.2	11.2	922	17	CNS0073W
4	101.8	11.1	919	17	CNS04ENV
5	100	10.9	1201	9	AL581589
6	99.4	10.8	1141	17	A0743305

C	7	99	10.8	1101	17	CNS00FXE	AL071370 Drosophila
C	8	96.6	10.5	1024	17	AG133080	AG133080 Pan trogl
C	9	95.8	10.4	1097	17	CNS01224	A1102202 Drosophila
C	10	95.8	10.4	1101	17	CNS0100X	A1038379 Drosophila
C	11	95	10.3	791	17	CNS009KS	AL053801 Drosophila
C	12	95	10.3	1101	17	CNS00LOO	AL053807 Drosophila
C	13	95	10.3	1144	17	A0743364	A0743364 HS_5387_B
C	14	94.4	10.3	1216	17	AG135357	AG135357 Pan trogl
C	15	94.2	10.2	815	17	B12686	B12686 F27J14-77 I
C	16	94	10.2	1188	17	AG135332	AG135332 Pan trogl
C	17	93.8	10.2	859	17	AG128925	AG128925 Pan trogl
C	18	93.8	10.2	1008	17	AG137085	AG137085 Pan trogl
C	19	93.8	10.2	1099	17	A0743360	A0743360 HS_5387_B
C	20	93.6	10.2	884	17	CNS006UO	AL065923 Drosophila
C	21	93.6	10.2	969	17	A0743309	A0743309 HS_5387_B
C	22	93.6	10.2	1081	17	AG135328	AG135328 Pan trogl
C	23	92.8	10.1	955	17	AG076494	AG076494 Pan trogl
C	24	92.8	10.1	981	17	AG127518	AG127518 Pan trogl
C	25	92.8	10.1	1353	17	AG128010	AG128010 Pan trogl
C	26	92.6	10.1	1183	17	AG136828	AG136828 Pan trogl
C	27	92.4	10.0	880	17	AG139490	AG139490 Pan trogl
C	28	92.2	10.0	885	17	CNS031TM	AL249807 Tetradon
C	29	92	10.0	1101	17	CNS00LIT	AL078714 Drosophila
C	30	92	10.0	1137	17	AQ743326	AQ743326 HS_5387_B
C	31	92	10.0	1225	17	CNS0166K	AL106358 Drosophila
C	32	92	10.0	1491	12	BE882936	BE882936 601505836
C	33	91.4	9.9	700	17	AG127423	AG127423 Pan trogl
C	34	91.4	9.9	866	17	AG126308	AG126308 Pan trogl
C	35	91.2	9.9	1054	17	AG137135	AG137135 Pan trogl
C	36	91	9.9	1749	13	BM415379	BM415379 OP20453 M
C	37	90.6	9.8	987	17	AG139684	AG139684 Pan trogl
C	38	90.6	9.8	927	17	CNS0142W	AL103634 Drosophila
C	39	90.2	9.8	1026	17	AG126150	AG126150 Pan trogl
C	40	90	9.8	953	17	CNS006PF	AL056567 Drosophila
C	41	90	9.8	1101	17	CNS017FC	AL107970 Drosophila
C	42	89.8	9.8	896	17	AG141027	AG141027 Pan trogl
C	43	89.8	9.8	1102	17	AG135459	AG135459 Pan trogl
C	44	89.4	9.7	909	17	CNS00JTL	AL076720 Drosophila
C	45	89.4	9.7	921	17	AG127701	AG127701 Pan trogl

#### ALIGNMENTS

RESULT 1  
CNS0071A/c 895 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
DEFINITION BACRI4B09 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL066286  
VERSION AL066286.1 GI:4945153

KEYWORDS Drosophila melanogaster.  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 895)  
Genoscope.

AUTHORS Direct Submission  
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BR 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osagawa and  
Aaron Mammoser in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial



Y	820	AGGAATCTAAGCCCTTCTCAAGTCTCCACACATTCGAGACCCATTCACACCTTCATC	879
Db	753	CAGCAGCCCAAAAAAACCCTCCACACAAAAAACAAGACACACACACACACACCA	812
Y	880	TCGACCCACACACGCGAGTATTTATTTAAAAAAA	919
Db	813	CCAAAACCCACAGAAAAACGACACGCGCCCAACACAAAAA	852
RESULT 3			
CNS0073W/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
<p>Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)</p> <p>Web : www.genoscope.cns.fr</p> <p>Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a>.</p> <p>Location/Qualifiers</p> <p>1..922</p> <p>/organism="Drosophila melanogaster"</p> <p>/db_xref="taxon:7227"</p> <p>/clone="BACR14D09"</p> <p>/clone_lib="RPCI-98"</p> <p>/note="end : TTT3"</p>			
BASE COUNT	223 a	95 c	109 g
ORIGIN			
Query Match	11.2%	Score 103.2;	DB 17; Length 922;
Best Local Similarity	24.8%	Pred. No. 1.9e-10;	
Matches 102;	Conservative 164;	Mismatches 144;	Indels 1; Gaps 1;
Y	386	CAACCCACACAGTCAAGACTATAAAACACACACACACCCCAACCCGACGAGCC	445
Db	921	MMMMCMCCMCCMCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM	862
Y	446	ACTCAAAAACAAGCGCAAAACCAACCAACCAACCAACCAATGATTTCACTTCGA	505
Db	861	ACAAAAMMMMMACMMMMCMACMMACMMACMMACMMACMMACMMACMMACMM	802
Y	506	GTGTTAACTTTGACCTCGACGATATGACGAACAATCAACCTGCTGGGCTATCTGC	565
Db	801	ACACM	743
Y	566	AAAAGATTCACCAAAAAACCGAAGAAAGAAACCAACCAACCAACGCTTCACAAAAACA	625

[illegible]

Matches	256: Conservative	39: Mismatches	264: Indels	4: Gaps	2
QY	337 AACCAACACCACTACTAGCTTCACACACACACAGAGATGACGTAAACCTGCAACCCACACAC	396			
Db	1152 AA	1093			
QY	397 AGTCAAGACTATAAAACACAAACAAACCCAAACACACACACACAGCCCTACTACAAACA	456			
Db	1092 AACCNACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1033			
QY	457 ACGCCAAAAACAACCAACCAACCAACCAATATATGATTTTTCACCTTGGAAGTGTAACTT	516			
Db	1032 ACCCCACAAAAAACAACCCCAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA	973			
QY	517 TGTACCTGACAGCATATGACGACAAATGCAACTGTGTGGCTATCTGTGAAAAAGATAC	576			
Db	972 AAACACAAACCCCAAAAAAAAAAACCACCAACCAACCAACCAACCAACCAACCAACCAAC	913			
QY	577 AAACAAAAAACCCGAGAAAGAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA	636			
Db	912 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	855			
QY	637 AACCAAAAAAGTCTCAAACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	696			
Db	854 AAC - AACACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	797			
QY	697 CACAGAGAAGCCAAACCATCAACACCCACCAACCAACCAACCAACCAACCAACCAACCA	756			
Db	796 AAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	737			
QY	757 CAACACACAGAGAAATCCAAAACCTACAAAGTCAAAATGAAACCTTTCACCTACACTCT	816			
Db	736 MAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	677			
QY	817 CGAAGGCAATCTAAAGCCCTTCTCAAGTCTGACACATCTCGAGACACCATCAACACCTC	876			
Db	676 CACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	617			
QY	877 ATCTCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	899			
Db	616 CACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	594			
RESULT 6					
A0743305					
LOCUS					
DEFINITION	A0743305	1141 bp	DNA	linear	GSS 16-JUL-1999
ACCESSION	HS-5387.B2.B02.SP6	RPCL-11	Human Male	BAC Library	Homo sapiens
VERSION	A0743305				
KEYWORDS	A0743305.1	GI:5520827			
SOURCE	GSS.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
	1 (bases 1 to 1141)				
	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,				
	Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and				
	Hood,L.				
TITLE	Sequence-tagged connectors: A sequence approach to mapping and				
JOURNAL	scanning the human genome				
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				
COMMENT	99380589				
	Contact: Mahairas GG, Wallace JC, Hood L				
	High Throughput Sequencing Center				
	University of Washington				
	401 Queen Anne Avenue North, Seattle, WA 98109, USA				
	Tel: (206) 616-3618				
	Fax: (206) 616-3887				
	Email: jwallace@u.washington.edu				
	Clones are derived from the human BAC library RPCL-11. For BAC				
	library availability, please contact Pieter de Jong				
	(pieter@dejong.med.buffalo.edu). Clones may be purchased from				

BACPAC Resources ([http://bacpac.med.bufileto.edu/ordering\\_bac.htm](http://bacpac.med.bufileto.edu/ordering_bac.htm))  
or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server:  
Plate: 963 row: D column: 4  
Seq primer: SP6

Class: BAC ends  
High quality sequence stop: 1141.

## FEATURES

source

Location/Qualifiers  
1..1141

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate:963 Col-4 Row-D"  
/clone\_lib="RPC1-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBACE3.6 vector at EcoRI sites"

BASE COUNT 760 a 301 c 16 g 15 t 49 others  
ORIGIN

Query Match 10.8%; Score 99.4; DB 17; Length 1141;

Best Local Similarity 46.7%; Pred. No. 1e-09; Mismatches 417; Indels 10; Gaps 4;

Matches 374; Conservative 0; Mismatches 417; Indels 10; Gaps 4;

120 TACGACAAATCATATTCATTCATTCGCAATGATATCTCATCTTCACTTATATTAATGACG 179

14 TNNNAAGANNAGACATNANNANCTNNNNNAGCTNNNCCCTCTNTAATTTNNAAGGGCGG 73

180 CCATCATATTCATAGGCTGGGCAACACAAATCATACACTACACTGCAATCATCATAG 239

74 CCAANNAAMCCCTTANAGAGAAANNTACANGACCNCNNNAANANANNAANNAANNA 133

240 ATCCACACAGCCGATGACAGACACACACCCCAATACCTGACTGAGATCTGACCTTG 299

134 AAACAAACAAAAATTAAAAACAAAAACAAAAACAAAAACAAAAACACACACACACAA 193

300 GAATCAGCTTCTCATATCTGTGAAATTAATCATCAACACACACACACATAGCTTCA 359

194 AACACA-----AAAACAAAAAACAAACACACACACACACACACAAAAACAAA 247

360 CAACACACGAGTCAAGTCAAACTGCAACCCACACACACTCAAGACTTAAACACACAA 419

248 CAAAAAACAA--CAAAAAACAAAAACAAAAACAAAAACAAAAACCAACACACACAA 305

420 CAACCCAAACACACCCACAGCCCATACAAACACACCCCAACCAACCAACCAACA 479

306 CAA 365

480 AACCCAAATGATTTTCACTTGAAGTGTTAACCTTGTACCTGACATATGACGCA 539

366 CAACAAANACAAACAAACACGACGACCAACACACACACACACACACACACAAAA 425

540 ACAATCAACCTGCTGGCTATCTGCAAAAGATACCAACAAAAACCAAGAGAAAA 599

426 ACAAAAAAACACACAAACACAAAAACAAAAACAAAAACCAACACACACACAA 485

600 CCACCAACCAAGCTCAAAAAACCAACCTTCAAGACACCAAAAAACCAAGTCTCAACCTC 659

486 CCAACACACACAAACACACAAACCAACACACACACACACACACACACACAA 544

660 AAACCACTAAACCAAGAGATGACACACACACACACACACACACACACACACACAA 719

545 CAACAAACCAAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 604

720 CCACCAACCAACCAACCACTACACCTGCTACCAACACACACACACAGAAATCCAAAC 779

605 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 664

780 TCACAAGTCAATGGAACCTTCCAGCTCACTCTCGAAGAGCAATTAAGCCCTTCTC 839

665 AAACACACAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 724

QY 840 AAGTCACACACATCCGACCCATCAACCTCATCTGCACCCACACACACGCC 899

DB 725 AAAAC-CAAAAAACACACACACACACACACACACACACACACACACACAC 783

QY 900 ACTAGTTATTTAAAAA 920

DB 784 AC 804

## RESULT 7

CNS00FHE/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Oseogawa and  
Aaron Mammoser in Pletier de Jong's laboratory in the department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPC1-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.bufileto.edu/drosophila\\_bac.htm](http://bacpac.med.bufileto.edu/drosophila_bac.htm).

## FEATURES

source

Location/Qualifiers  
1..1101

/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACR32C19"  
/clone\_lib="RPC1-98"  
/note="end : TET3"

BASE COUNT 326 a 104 c 183 g 186 t 302 others  
ORIGIN

Query Match 10.8%; Score 99; DB 17; Length 1101;

Best Local Similarity 25.6%; Pred. No. 1.3e-09; Mismatches 157; Indels 0; Gaps 0;

Matches 104; Conservative 145; Mismatches 157; Indels 0; Gaps 0;

QY 358 AACACACGAGTCAAGTCAACCTGCAACCCACACAGTCAAGCTAAACACACAC 417

DB 1081 AAAC 1022

QY 418 AAC 477

DB 1021 AAAC 962

QY 478 CAAACCAATTAATGATTTCACTTGAAGTGTTAACCTTGTACCTGACATATGACG 537

DB 961 MAAAC 902

QY 538 CAACATTCACACCTGCTGAGTATCTGCAACCAACCAACCAACCAACCAACCA 597







LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
CNS009KS	791 bp DNA linear	GSS 03-JUN-1999								
BACR19F04	Drosophila melanogaster genome survey sequence T7 end of BAC # 11Y), genomic survey sequence.	AL053801								
AL053801	AT053801.1 GI:4935176									
GSS.	Drosophila melanogaster.									
Drosophila melanogaster.	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryygota; Nephelera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.									
Genoscope.	1 (bases 1 to 791)									
Direct Submission	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage									
BP 191 91006 EVRI cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr										
- Web : www.genoscope.cns.fr)										
Determination of this BAC-end sequence was carried out as part of a										
collaboration with the Berkeley Drosophila Genome Project (BDGP).										
The BDGP is constructing a physical map of the Drosophila										
melanogaster genome using these BACs. For further information										
please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila										
melanogaster BAC library was prepared by Kazuhiro Oseegawa and										
Aaron Mamosser in Pieter de Jong's laboratory in the Department of										
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,										
NY. The library is named RPCI-98 and was constructed by partial										
ECORI digestion of Drosophila DNA provided by the BDGP from the										
isogenic strain Y2: cn bw sp, the same strain used for the BDGP's										
PI and EST libraries. A more detailed description of the library										
and how to order individual BAC clones, the entire library, or										
filters for hybridization from the BACPAC Resource Center can be										
found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .										
Location/Qualifiers										
1..791										
/organism="Drosophila melanogaster"										
/db_xref="taxon:7227"										
/clone="BACR19F04"										
/clone_lib="RPCI-98"										
/note="end : 17"										
BASE COUNT	464 a 105 c 67 g 78 t 77 others									
ORIGIN										
Query Match	10.3%; Score 95; DB 17; Length 791;									
Best Local Similarity	47.1%; Pred. No. 7.7e-09;									
Matches 205; Conservative 43; Mismatches 178; Indels 9; Gaps 2										
358 AACCAACACGAGGAGTCAGTCTGCAACCTGCAAGCCCAAGTCAGATCAAAACGACAC	417									
DB 33 AAAAAAAAAAAMCMCAAAACCAACCAAAAAAAAAAACAAGCAAAAGAAAAACAA	92									
418 AACCAACCCAAACACCAAGCCCAAGCCCAACCAACCAACCAACCAACCAACCA	477									
DB 93 AAAAAAAAAAAMCAAAACCAACCAAAAAAAAAAACAAGCAAAAGAAAAACAA	152									
478 CAACCAACCA-ATATGATTTTCACTTCTGGAAGTGTTAACCTTGAACCTGCACATATGCA	536									
DB 153 AAAAAAAAAAAMCAAAACCAACCAAAAAAAAAAACAAGCAAAAGAAAAACAA	212									
537 GCAACATTCACACTGCTGGGCTATCTGCAAAAGATATACCAACCAAAACCAAGAAAGA	596									
DB 213 ACWATTAAGCAACAA-----AAAMAMMAAAAAAAAAAAAAAAAAAAAAAAAAA	264									
597 AAACCAACCAAGCCTGCTGCAAAACCAACCAACCAACCAACCAACCAACCAACCA	656									
DB 265 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	324									
657 CTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	716									

[illegible]

```

Db      695  MTTAATTCACIMMAMAMAMITMTTMMAMACSRMAMIMMCCSCSAMCMAMAMAMMAMMAMMA 754
Oy      609  AGCCATCAAAAAAAGCCACCTTCAAGACACCAAAAAAGATCTCAAACTCAACCACTA 668
Db      755  WIMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 814
Oy      669  AACCAAGAGTACCCACCCACCAAGCCACAGAAAGCCACATCAACSCCAAAAA 728
Db      815  YACSCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 874
Oy      729  CAACATTCACACTACACTGTCTACCAACCAACACAGCAAAATCCAAACTCAAGTC 788
Db      875  CCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 934
Oy      789  AATGGAACCTTCACATCAACCTTCACAGGACATCTAAGCCCTTCACACTTCA 848
Db      935  MAMMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 994
Oy      849  CAACATCCGAGCACCACATCAACACCTTCATCTCCACCAACACACAGCCAGTA 903
Db      995  MAMMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 1049

```

```

RESULT 13
LOCUS   A0743364               1144 bp      DNA      linear      GSS 16-JUL-1999
DEFINITION
HS:5387.B2.H09.SP6.RPCI-11 Human Male BAC Library Homo sapiens
ACCESSION
A0743364
VERSION
A0743364.1 GI:5520886
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (bases 1 to 1144)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

```

```

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pjeter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.hsc.washington.edu
Plate: 963 row: P column: 18
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 1144.
Location/Qualifiers
1..1144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lib="Plate:963 Col=18 Row=P"
/clone_1lib="RPCI-11 Human Male BAC Library"
/sex="male"

```

```

FEATURES
source
PAC3.6 vector at EcoRI sites"

```

```

BASE COUNT      625 a      443 c      11 g      11 t      54 others
ORIGIN

```

```

Query Match      10.3%; Score 95; DB 17; Length 1144;
Best Local Similarity 46.7%; Pred No. 7.5e-09;
Matches 371; Conservative 0; Mismatches 420; Indels 3; Gaps 3;

```

```

Oy      125  CAATCATATTATTCATTTCTGGCAATGATATCTCACTTCACTTATTAATCAGCATC 184
Db      340  CAAACAAACACACCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 399
Oy      185  ATTTGATGCTCTGGGCAACGCAAAAGTACAGCTTAAGCACTGATCTTAAAGATGCA 244
Db      400  AAAAAAANAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 459
Oy      245  ACAAGCAGATCAAGAACAACCAACCCCA-CATACCTCATCTGAGATCTCTGCTTGAT 303
Db      460  CCACACACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 519
Oy      304  CAGCTTCTCAATCTGTCTGAAATTTACATCAACCAACCAACCAACCAACCAACCAAC 363
Db      520  CACACAAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 579
Oy      364  ACCAGAGTCAGTCAAACTGCAACCTGCAACCCCAACAGTCAGATTAAGTAACCAACCA 423
Db      580  CCACAAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 638
Oy      424  CCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 483
Db      639  ACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 698
Oy      484  CAATATGATTTTCACTTCTGCAAGTGTGTTTACCTTTGACCTGACATATGACGACAA 543
Db      699  CAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 757
Oy      544  TCCAACTGCTGGGCTATCTGCAAAAGATTAACCAACCAACCAACCAACCAACCAACCA 603
Db      758  CAAAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 817
Oy      604  CACCAAGCTTCAAAAAAAGCACTTCAAGCAACCAACCAACCAACCAACCAACCAACCA 663
Db      818  AACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 877
Oy      664  CACTAAACCAAGAGTATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 723
Db      878  ACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 937
Oy      724  CAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 783
Db      938  CACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 997
Oy      784  AAGTCAAAATGGAACCTTCCTCACTCACTCTCCGAAGCAATTAAGCCCTTCTCAAGT 843
Db      998  CACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1057
Oy      844  CTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 903
Db      1058  AAAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1117
Oy      904  GTTATTAAAAAAA 917
Db      1118  CCAAAACCAACCAAAA 1131

```

```

RESULT 14
LOCUS   AG135357               1216 bp      DNA      linear      GSS 04-NOV-2001
DEFINITION
Pan troglodytes DNA, clone: PTB-148J19.F, genomic survey sequence.
ACCESSION
AG135357
VERSION
AG135357.1 GI:16665035
KEYWORDS
GSS.
Pan troglodytes male lymphoblast DNA, clone_1lib:PTB Chimpanzee Male
BAC library clone:PTB-148J19.F.

```

ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE 1  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1216)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. 1216  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-148J19.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_11b="PTB Chimpanzee Male BAC Library"  
BASE COUNT 572 a 532 c 15 g 29 t 68 others  
ORIGIN

Query Match 10.3%; Score 94.4; DB 17; Length 1216;  
Best Local Similarity 47.1%; Pred. No. 9.9e-09;  
Matches 319; Conservative 0; Mismatches 353; Indels 6; Gaps 2;

222 CAACTGCAATCATACAGATGACAGACAGGAGATCAAGACCAACCCCAATACCTCA 281  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 109 CAAAANNNAAACNNCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAN 168  
282 CTCGAGATCTCAGCTTGAATCAGCTTCTCAATCTGTGAAATTAATCAACAACA 341  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 169 AANAANCCACACCAAAATTCACCCACAAACACCCCAACCCCAACCAACAACA 228  
342 CCACCACTAGCTTCAACACACACAGAGTCAAGTCAACCTGCAACCCCAACAGTCA 401  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 229 AACCCAGACACACACACCCCAACCAACCCCAACCCCAACCCCAACCAACA 288  
402 AGACTAAAAACACACACACCAACCAACCCCAACCCCAACCCCAACCAACAGCC 461  
289 CACCCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 348  
462 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 521  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 349 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 403  
522 CCTGCAAGCATATGACCAACATCACTGTGGCTATCTGCAAAAGAAATACCAACA 581  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 404 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 463  
582 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 641  
464 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 523  
642 AAAAAAGATCTCAACCTCAACCAACCAACCAACCAACCAACCAACCAACCAAC 701  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 524 AAAAAACCAACCAACCAACCTCTGCTACCAACCAACCAACCAACCAACCAACCA 583

QY 702 AAGAGCAACCAATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 761  
Db 584 CACAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 642  
QY 762 CCACAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 821  
Db 643 CCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 702  
QY 822 GCAATCTAAGCCTTCTCAAGTCTCCACAACATCCGAGCACCCTACACCTCATCTC 881  
Db 703 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 762  
QY 882 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 899  
Db 763 CACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780

RESULT 15  
B12686/c 815 bp DNA linear GSS 14-MAY-1997  
LOCUS  
DEFINITION F27J14-T7 IGF Arabidopsis thaliana genomic clone F27J14, DNA  
sequence.  
ACCESSION B12686  
VERSION B12686.1 GI:2093806  
KEYWORDS GSS.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 815)  
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and  
Ecker, J.  
BAC End Sequences at ATGC  
JOURNAL Unpublished (1997)  
CONTACT Ecker, J  
ORGANISM Arabidopsis thaliana  
University of Pennsylvania  
Dept. of Biology, University of Pennsylvania, Philadelphia, PA  
19104  
Tel: 215-898-9384  
Fax: 215-898-8780  
Email: jecker@atgenom.bio.upenn.edu  
Seq primer: T7  
Class: BAC ends  
High quality sequence start: 400  
High quality sequence stop: 417.  
Location/Qualifiers  
1. 815  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="F27J14"  
/clone\_11b="IGF"  
/sex="hermaphrodite"  
/note="vector: BelobacII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altmann"  
BASE COUNT 22 a 11 c 185 g 502 t 95 others  
ORIGIN

Query Match 10.2%; Score 94.2; DB 17; Length 815;  
Best Local Similarity 45.0%; Pred. No. 1.1e-08;  
Matches 297; Conservative 0; Mismatches 355; Indels 8; Gaps 1;

220 AACCACTGCAATCATACAGATGCAACAGGAGATCAAGACCAACCCCAATACCT 279  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 812 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753  
QY 280 CACTAGATCTCAGCTTGAATCAGCTTCTCAATCTGTGAAATTAATCAACAAC 339  
Db 752 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693  
QY 340 CACCACTAGCTTCAACAGACACAGAGTCAAGTCAACCTGCAACCAACCAACAGT 399

Db	692	AAACGACAAACCGCAAAAAAAAAAAAAAAAAACGACACAAAAAAAAAAAAAAAAACGACAA	633
QY	400	CAAGACTAAAAACACACACACACCCAAACGAAACCGAGAGCCCACTACAAACAGC	459
Db	632	AAACGACAAAAAAACACACACACCCCAAAAAAACACACACAAAAAACACAAACAGCA	573
QY	460	CCAAACGAAACCCCAAAACCAACCCCAATATGATTCTTCACTTGAAGTGTTTAACTTGT	519
Db	572	AAAAAAAAACCAACCCAAAAANANANAAAAACAAAAACCAACCAAAAAAAAAAAAAAN	513
QY	520	ACCGTCGAGCATATGACGACAACTCCAACTGCTGGGCTATCTGCAAAAGAAATACCAA	579
Db	512	AAAAACCAAAACACAAACAAAAACCAACAAAAACAAAAACAAACACACAAAAAACAC	453
QY	580	CAAAAAACCGAGAAAGAAACCCACCCACCAAGCTACAAAAAAACCAACCTTCAAGACA	639
Db	452	AAAAAAAAACCAAAAAANCAANAAACCCCNCCNCAAAAAAAANNNNNNN	393
QY	640	CAAAAAAGATCTCAACCTCAACCA-----CTTAACCAAGGAAGTACCCACACC	691
Db	392	NNNCACANANCCCAANNNACACACAAAAAAACAAAAACAAANNCAAAAACCAAAACA	333
QY	692	AAGCCACAGGAAGCCACCATCAACACCCAAAAACAACATCACTACACTGCTC	751
Db	332	AAACCCAAACCAAAACCCACCAACAAAAAAACAAAAAAACAAACCCCAACAAACACA	273
QY	752	ACGACAAACACACAGGAATCCAAAACCTACAAGTCAAAATGGAACCTTCCACTACCC	811
Db	272	ACAAAAAANANCAAAAAAANNNNANNAACANNNCCAAACAAACACAAAAANNAAC	213
QY	812	TTCCTCGAAGCAATCTTAAGCCCTTCTCAAGTCTCCGACAAACATCGAGACCACTACAA	871
Db	212	CCGACACCAACCAAAAAAACCAAAAAACCAACCCACACCCCCCAACAAAAA	153

Search completed: May 1, 2003, 12:33:26  
Job time : 1146.76 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 12:33:38 ; Search time 37.6717 Seconds  
(Without alignments)  
1054.072 Million cell updates/sec

Title: US-09-462-816-2  
Perfect score: 1544  
Sequence: 1 MSKNKDQRTAKLEKTWDTL.....VSTTSEHPQSSPPTTRQ 298

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544	100.0	298	14	AA39286
2	1544	100.0	298	20	AA96313
3	1418	91.8	298	8	AA70845
4	1418	91.8	298	13	AA25302
5	1418	91.8	298	19	AA47605
6	1418	91.8	298	23	AAU74676
7	1229	79.6	232	20	AAW6314
8	901	58.4	681	10	AA90441
9	771	49.9	299	22	AA68336
10	572.5	37.1	452	22	AA67771

11	562	36.4	349	17	AA95660
12	562	36.4	349	22	AA68028
13	558	36.1	349	16	AA88253
14	558	36.1	101	17	AA95610
15	558	36.1	101	17	AA95616
16	558	36.1	101	17	AA97050
17	558	36.1	101	20	AA44078
18	558	36.1	101	21	AA18805
19	558	36.1	101	22	AA67741
20	558	36.1	101	22	AA84123
21	558	36.1	101	22	AA68016
22	558	36.1	101	22	AA67775
23	552.5	35.8	356	17	AA95661
24	538	34.8	101	16	AA88255
25	538	34.8	101	17	AA95612
26	538	34.8	101	17	AA95618
27	538	34.8	101	17	AA97052
28	538	34.8	101	17	AA97051
29	538	34.8	101	20	AA44080
30	538	34.8	101	22	AA97311
31	534	34.6	101	20	AA97310
32	514	33.3	101	20	AA97312
33	506	32.8	101	17	AA95614
34	506	32.8	101	17	AA97063
35	506	32.8	101	20	AA44090
36	506	32.8	101	22	AA84135
37	470.5	30.5	92	17	AA95615
38	353	22.9	61	17	AA97072
39	353	22.9	61	20	AA44099
40	353	22.9	61	22	AA84144
41	343	22.2	59	17	AA97073
42	343	22.2	59	20	AA44100
43	343	22.2	59	22	AA84145
44	333	21.6	57	17	AA97074
45	333	21.6	57	20	AA44101

## ALIGNMENTS

RESULT 1	ID	AA39286	standard; Protein; 298 AA.
XX	AC	AA39286;	
XX	DT	13-JAN-1994	(first entry)
XX	DE	Respiratory syncytial virus (RSV) G protein.	
XX	KW	PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.	
XX	OS	Respiratory syncytial virus.	
XX	PN	W09314207-A.	
XX	PI	22-JUL-1993.	
XX	PD	05-JAN-1993;	93WC-CA00001.
XX	PF	06-JAN-1992;	92GB-0000117.
XX	PA	(CONN-) CONNAUGHT LAB LTD.	
XX	PI	Ewasysghyn ME, Klein MH;	
XX	DR	WPI, 1993-243222/30.	
XX	DR	N-PSDB; AAQ45686.	
XX	PT	Multimeric hybrid genes and their chimeric proteins - are	
XX	PT	vaccines against multiple pathogenic infections e.g.	
XX	PT	para-influenza virus and respiratory syncytial virus	

Streptococcal prot  
Anti: o acid sequenc  
RSV subgroup A wil  
RSV subgroup. A prote  
RSV subgroup A wi  
Respiratory Syncyt  
RSV G protein anti  
A G2Na peptide der  
Amino acid sequenc  
Amino acid sequenc  
Amino acid sequenc  
Amino acid sequenc  
Immunogenic carrie  
RSV subgroup A mod  
RSV subgroup. A prote  
RSV subgroup A mu  
Respiratory Syncyt  
RSV G protein anti  
Peptide which indu  
Peptide which indu  
Peptide which indu  
RSV subgroup. A prote  
Respiratory Syncyt  
RSV G protein anti  
Amino acid sequenc  
Respiratory Syncyt  
RSV G protein anti  
Amino acid sequenc  
Respiratory Syncyt  
RSV G protein anti

PS Claim 11; Figure 7A-7D; 80pp; English.  
XX  
CC A novel multimeric hybrid gene is used as a vaccine. The gene  
CC consists of two gene sequences which are linked and encode antigenic  
CC regions, these two sequences being derived from two different  
CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
CC (RSV)). The gene sequences that are particularly used are those  
CC which encode PIV-3 F and HN proteins (AA045683, AA045684) and RSV F and  
CC G proteins (AA045685, AA045686).  
XX  
SQ

Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 14; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.1e-113;

Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60  
DB 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60  
QY 61 FIASANHKVLTTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSEITSOITTLASTP 120  
DB 61 FIASANHKVLTTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSEITSOITTLASTP 120  
QY 121 GYKSNLOPTTVKTKNTTTOPTOPSKPTTKORONKPPKNPNDFEFVFNFPVCSICSNP 180  
DB 121 GYKSNLOPTTVKTKNTTTOPTOPSKPTTKORONKPPKNPNDFEFVFNFPVCSICSNP 180  
QY 181 TCAICKRIIPNKKRGGKTTTPKPKPTTKKDLKPOTTKPEVPTTKPEPTINTTK 240  
DB 181 TCAICKRIIPNKKRGGKTTTPKPKPTTKKDLKPOTTKPEVPTTKPEPTINTTK 240  
QY 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298  
DB 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

#### RESULT 2

AAW96313  
ID AAW96313 standard; Protein; 298 AA.

XX  
AC AAW96313;

DT 28-JUN-1999 (first entry)

DE Membrane bound G protein of respiratory syncytial virus.

XX  
KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

XX  
KW tissue plasminogen activator.

OS Respiratory syncytial virus.

XX  
PN WO9904010-A1.

XX  
PD 28-JAN-1999.

XX  
PF 16-JUL-1998; 98WO-CA00697.

XX  
PR 18-JUL-1997; 97US-0896442.

XX  
PA (CONN-) CONNAGHT LAB LTD.

XX  
PI Klein MH, Li X, Sambhara S;

XX  
DR WPI: 1999-132254/11.

XX  
DR N-PSDB; AAX08421.

PT Immunogenic composition for generating antibodies against  
PT respiratory syncytial virus - comprises non-replicating vector  
PT containing the protein G sequence, useful in protective vaccines and  
PT to raise antibodies for diagnosis  
XX

PS Claim 4; Fig 2; 67pp; English.

XX  
CC The respiratory syncytial virus (RSV) G protein can be used in  
CC vaccines by inserting the G protein gene into a non-replicating  
CC vector. The G protein is placed under the control of alternative  
CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G protein's  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.  
XX

SQ Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 20; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.1e-113;

Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60  
DB 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60  
QY 61 FIASANHKVLTTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSEITSOITTLASTP 120  
DB 61 FIASANHKVLTTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSEITSOITTLASTP 120  
QY 121 GYKSNLOPTTVKTKNTTTOPTOPSKPTTKORONKPPKNPNDFEFVFNFPVCSICSNP 180  
DB 121 GYKSNLOPTTVKTKNTTTOPTOPSKPTTKORONKPPKNPNDFEFVFNFPVCSICSNP 180  
QY 181 TCAICKRIIPNKKRGGKTTTPKPKPTTKKDLKPOTTKPEVPTTKPEPTINTTK 240  
DB 181 TCAICKRIIPNKKRGGKTTTPKPKPTTKKDLKPOTTKPEVPTTKPEPTINTTK 240  
QY 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298  
DB 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

#### RESULT 3

AAP70845  
ID AAP70845 standard; protein; 298 AA.

XX  
AC AAP70845;

DT 05-APR-1991 (first entry)

DE Sequence of human respiratory syncytial virus (HRSV) A2 strain

XX  
DE G protein.

XX  
KW Vaccine.

XX  
OS Human respiratory syncytial virus (HRSV).

XX  
FN WO8704185-A.

XX  
PD 16-JUL-1987.

XX  
PF 23-DEC-1986; 86WO-US02756.

XX  
PR 14-JAN-1986; 86US-0818740.

XX  
PA (UNIC-) UNIV OF N CAROLINA.

XX  
PA (WERTZ) WERTZ G W.

XX  
PA (WERTZ) WERTZ G W.  
XX  
DR WPI: 1987-206300/29.  
DR N-PSDB; AAN70784.  
XX



PT Vaccines for human respiratory virus - comprising proteins or  
PT fragment encoded by a DNA sequence coding for human respiratory  
PT syncytial virus proteins.  
XX  
PS Disclosure; Chart 13; 57pp; English.  
XX  
CC A novel plasmid which comprises a DNA sequence encoding this  
CC protein, and the protein itself, are claimed, for use as HRSV  
CC vaccines. The vaccine can be administered to pregnant women or to  
CC women of child bearing age to stimulate maternal antibodies.  
CC Infants can also be vaccinated at 2-3 months of age.  
XX  
SQ Sequence 298 AA;  
Query Match 91.8%; Score 1418; DB 8; Length 298;  
Best Local Similarity 93.3%; Pred. No. 1.6e-103;  
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
QY 1 MSKKNKQRTAKTLEKTDWTLNHLLEFISGLYKLNLSVAQITLSILAMISTSLITTAII 60  
DB 1 MSKKNKQRTAKTLEKTDWTLNHLLEFISGLYKLNLSVAQITLSILAMISTSLITTAII 60  
QY 61 FIASANHKVTLTALIIDATSOIKNTPTVLTQDPOLGISFNSLSEITTSQTTTILASTTP 120  
DB 61 FIASANHKVPTTALIIDATSOIKNTPTVLTQDPOLGISFNSLSEITTSQTTTILASTTP 120  
QY 121 GVKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPNNDHFHFVFNVPSCISNNP 180  
DB 121 GVKSTLOSTTVKTKNTTTQTOPSKPTTKORONKPPNNDHFHFVFNVPSCISNNP 180  
QY 181 TCWAIKRIIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTKPTPEEPTINTTK 240  
DB 181 TCWAIKRIIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTKPTPEEPTINTTK 240  
QY 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPOSVSTSEHPSPSPPTTRQ 298  
DB 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPOSVSTSEHPSPSPPTTRQ 298  
RESULT 4  
AAR25302  
ID AAR25302 standard; Protein; 298 AA.  
XX  
AC AAR25302;  
XX  
DT 03-MAR-1993 (first entry)  
XX  
DE HRSV glycoprotein G (gpc).  
XX  
KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;  
KW major capsid protein; N.  
XX  
OS Human respiratory syncytial virus strain A2.  
XX  
PN US5149650-A.  
XX  
PD 22-SEP-1992.  
XX  
PF 14-JAN-1986; 86US-0818740.  
XX  
PR 14-JAN-1986; 86US-0818740.  
XX  
PR 13-JUL-1988; 88US-0218737.  
XX  
PA (UYNC-) UNIV NORTH CAROLINA.  
XX  
PI Collins PL, Wertz GW;  
XX  
DR WPI: 1992-340247/41.  
XX  
DR N-PSDB; AAO29623.  
XX  
PT Vaccines for human respiratory virus - include structural genes  
PT coding for native structural viral proteins and immunogenic  
PT fragments

XX  
PS Disclosure; Page 18; 21pp; English.  
XX  
CC The sequences of mRNA encoding HRSV structural proteins are given in  
CC AAO29622-26. The proteins are F, G, 22K, 9.5K and major capsid  
CC protein N. The sequences and encoded proteins are useful for  
CC preparing vaccines against HRSV. The vaccines can be used to confer  
CC immunity against respiratory tract infections on human subjects.  
XX  
SQ Sequence 298 AA;  
Query Match 91.8%; Score 1418; DB 13; Length 298;  
Best Local Similarity 93.3%; Pred. No. 1.6e-103;  
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
QY 1 MSKKNKQRTAKTLEKTDWTLNHLLEFISGLYKLNLSVAQITLSILAMISTSLITTAII 60  
DB 1 MSKKNKQRTAKTLEKTDWTLNHLLEFISGLYKLNLSVAQITLSILAMISTSLITTAII 60  
QY 61 FIASANHKVTLTALIIDATSOIKNTPTVLTQDPOLGISFNSLSEITTSQTTTILASTTP 120  
DB 61 FIASANHKVPTTALIIDATSOIKNTPTVLTQDPOLGISFNSLSEITTSQTTTILASTTP 120  
QY 121 GVKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPNNDHFHFVFNVPSCISNNP 180  
DB 121 GVKSTLOSTTVKTKNTTTQTOPSKPTTKORONKPPNNDHFHFVFNVPSCISNNP 180  
QY 181 TCWAIKRIIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTKPTPEEPTINTTK 240  
DB 181 TCWAIKRIIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTKPTPEEPTINTTK 240  
QY 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPOSVSTSEHPSPSPPTTRQ 298  
DB 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPOSVSTSEHPSPSPPTTRQ 298  
RESULT 5  
AAM47605  
ID AAM47605 standard; Protein; 298 AA.  
XX  
AC AAM47605;  
XX  
DT 11-JUN-1998 (first entry)  
XX  
DE HRSV glycoprotein G.  
XX  
KW HRSV; glycoprotein F; gpf; glycoprotein G; gpc; vaccine.  
XX  
OS Human respiratory syncytial virus.  
XX  
PN US5716823-A.  
XX  
PD 10-FEB-1998.  
XX  
PF 12-MAY-1997; 97US-0854783.  
XX  
PR 13-JUL-1988; 88US-0218737.  
XX  
PR 14-JAN-1986; 86US-0818740.  
XX  
PR 23-DEC-1986; 86MO-US02756.  
XX  
PR 11-JUN-1992; 92US-0887171.  
XX  
PR 12-MAY-1997; 97US-0854783.  
XX  
PA (PHAA) PHARMACIA & UPJOHN CO.  
XX  
PI Collins PL, Wertz GW;  
XX  
DR WPI: 1998-144802/13.  
XX  
DR N-PSDB; AAV18736.  
XX  
PT Production of human respiratory syncytial virus glyco-protein F or G  
PT - by culturing eukaryotic host cells transfected with corresponding  
PT DNA

PS Example 1; Columns 27-28; 17pp; English.

XX The present sequence was used in the development of a novel method

CC for the production of human respiratory syncytial virus (HRSV)

CC glycoprotein F (gpf) or glycoprotein G (gpc). The method comprises

CC culturing eukaryotic host cells transfected with an isolated DNA

CC sequence encoding HRSV gpf or gpc. The gp can be used to prepare

CC vaccines against HRSV.

XX

SQ Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 19; Length 298;

Best Local Similarity 93.3%; Pred. No. 1.6e-103;

Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKTDPTLNHLFISSGLYKLNKSAVOITSLIAMIISTSLITAAII 60

DB 1 MSKNDORTAKTLEKTDPTLNHLFISSGLYKLNKSAVOITSLIAMIISTSLITAAII 60

QY 61 FIASANKVLTFTAIIDATSOIKNTTPYLYLTOPOLGISFSNISEITSOITTLASTTP 120

DB 61 FIASANKVLTFTAIIDATSOIKNTTPYLYLTOPOLGISFSNISEITSOITTLASTTP 120

QY 121 GYKSNLOPTTVKTKNTTTOTOPSKPTTKORONKPPKPNDFHEFVFNVPSCISNNP 180

DB 121 GYKSTLOSTTVTKNTTTOTOPSKPTTKORONKPPKPNDFHEFVFNVPSCISNNP 180

QY 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

DB 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

QY 241 TTTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

DB 241 TTTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 6

AAU74676

ID AAU74676 standard; Protein; 298 AA.

XX

AC AAU74676;

XX

DT 09-APR-2002 (first entry)

XX

DE Respiratory syncytial virus G protein.

XX

KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;

KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;

KW antiviral chemotherapeutic compound; humoral response;

KW cellular immune response; hpiV; paediatric respiratory disease;

KW globin gene transfer; sickle cell disease; beta-thalassaemia;

KW human immunodeficiency virus infection; HIV.

XX

OS Human respiratory syncytial virus.

XX

PN WO200192548-A2.

XX

PD 06-DEC-2001.

XX

PF 22-MAY-2001; 2001WO-US16610.

XX

PR 01-JUN-2000; 2000US-208701P.

XX

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX

PI Portner A, Takimoto T;

XX

DR MPI; 2002-130534/17.

XX

DR N-PSDB; AAS21045.

XX

PT Recombinant Sendai virus useful in vaccines to protect infection by

PT paramyxoviruses, comprises exogenous nucleic acid encoding

PT paramyxovirus protein or its antigenic fragment

XX

PS Disclosure; Page 48; 57pp; English.

XX

CC The invention relates to a recombinant Sendai virus comprising an

CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its

CC antigenic fragment. The virus may be administered in combination

CC with an antiviral chemotherapeutic compound. Two or more viruses

CC expressing different PMV proteins may be co-administered. Compositions

CC comprising the virus are useful for eliciting a humoral and/or

CC cellular immune response to a PMV in a mammal, particularly a human.

CC Further a recombinant Sendai virus comprising an exogenous nucleic acid

CC encoding a second PMV protein is also administered and priming and/or

CC boosting humoral or cellular immune response comprises administering

CC one or more of a recombinant or isolated PMV protein or its antigenic

CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral

CC vector encoding a PMV protein. The recombinant virus is useful as an

CC effective vaccine against hpiV or RSV (the major causes of paediatric

CC respiratory disease) and also to express any gene of

CC interest in target cells, providing a positive medical impact on

CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)

CC into stem cells effects a cure for sickle cell disease or beta-

CC thalassaemia. The recombinant virus may also prove effective in

CC conferring immunity to human immunodeficiency virus (HIV) infection.

CC The Sendai virus replicates at level that is high enough to

CC induce sufficient immunity, but does not cause any harm to human

CC recipient. The present sequence represents a respiratory syncytial

CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein

CC suitable for expression by the recombinant virus of the invention.

XX

SQ Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 23; Length 298;

Best Local Similarity 93.3%; Pred. No. 1.6e-103;

Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKTDPTLNHLFISSGLYKLNKSAVOITSLIAMIISTSLITAAII 60

DB 1 MSKNDORTAKTLEKTDPTLNHLFISSGLYKLNKSAVOITSLIAMIISTSLITAAII 60

QY 61 FIASANKVLTFTAIIDATSOIKNTTPYLYLTOPOLGISFSNISEITSOITTLASTTP 120

DB 61 FIASANKVLTFTAIIDATSOIKNTTPYLYLTOPOLGISFSNISEITSOITTLASTTP 120

QY 121 GYKSNLOPTTVKTKNTTTOTOPSKPTTKORONKPPKPNDFHEFVFNVPSCISNNP 180

DB 121 GYKSTLOSTTVTKNTTTOTOPSKPTTKORONKPPKPNDFHEFVFNVPSCISNNP 180

QY 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

DB 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

QY 241 TTTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

DB 241 TTTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 7

AAW96314

ID AAW96314 standard; Protein; 232 AA.

XX

AC AAW96314;

XX

DT 28-JUN-1999 (first entry)

XX

DE Secreted G protein of respiratory syncytial virus.

XX

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

KW tissue plasminogen activator.

XX

OS Respiratory syncytial virus.

XX

PN WO9904010-A1.

```

XX 28-JAN-1999.
PD 16-JUL-1998; 98WO-CA00697.
XX 18-JUL-1997; 97US-0896442.
XX (CONN-) CONNAUGHT LAB LTD.
XX Klein MH, Li X, Sambhara S;
XX WPI: 1999-132254/11.
DR N-PSDB: AAX08422.
XX
PT Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis
XX
PS Claim 9; Fig 3; 67pp; English.
XX
CC The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternative
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods. This truncated G
CC protein is secreted since it lacks a transmembrane domain.
XX
SQ Sequence 232 AA:
XX
Query Match 79.6%; Score 1229; DB 20; Length 232;
Best Local Similarity 100.0%; Pred. No. 8.3e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 HKVTLTTAIIODATSOIKNTPTVLTQDPQIGISFSNLSSEITSGTTIIILASTTPGVKSNL 126
DB 1 HKVTLTTAIIODATSOIKNTPTVLTQDPQIGISFSNLSSEITSGTTIIILASTTPGVKSNL 60
QY 127 QPTVTKTKNTTTOTOPSPKPTTKORONKPPNNDPHEFVNFVPCSCSNPTCWAIC 186
DB 61 QPTVTKTKNTTTOTOPSPKPTTKORONKPPNNDPHEFVNFVPCSCSNPTCWAIC 120
QY 187 KRIIPNKKPGKKTTPKPTKPTTKKDKLPQTKPKKEVPTTKPEEPTINTTKNTITTT 246
DB 121 KRIIPNKKPGKKTTPKPTKPTTKKDKLPQTKPKKEVPTTKPEEPTINTTKNTITTT 180
QY 247 LLTNNTGNPKLTSQMEFHTSSSEGNLSPSOVSTTSEHPSSSPPTTRQ 298
DB 181 LLTNNTGNPKLTSQMEFHTSSSEGNLSPSOVSTTSEHPSSSPPTTRQ 232
XX
RESULT 8
AAP90441
ID AAP90441 standard; protein: 681 AA.
XX
AC AAP90441;
XX
DT 01-NOV-1989 (first entry)
XX
DE Chimeric human respiratory syncytial virus glycoproteins F and G.
XX
KW Chimeric polypeptide: human respiratory syncytial virus;
XX protein F; protein G; vaccine.
XX
OS Human respiratory syncytial virus.

```

```

XX MO8905823-A.
PN 29-JUN-1989.
XX
XX 31-OCT-1988; 88WO-US03784.
XX 23-DEC-1987; 87US-0137387.
XX
XX (UPJO ) UPJOHN CO.
XX
XX Mathen M;
XX
DR WPI: 1989-206593/28.
XX
PT Chimeric human respiratory syncytial virus polypeptides(s)
PT - contg. immunogenic fragments from HNSV glycoproteins
PT F and G, for vaccine prodn.
XX
PS Claim 3; page 47-48; 50pp; English.
XX
CC Chimeric polypeptide contg. a signal sequence and one or more
CC immunogenic fragments from both human respiratory syncytial virus
CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg
CC E. coli, Chinese hamster ovary cells, murine C127 cells and
CC S. frugiperda.
XX
SQ Sequence 681 AA:
XX
Query Match 58.4%; Score 901; DB 10; Length 681;
Best Local Similarity 91.0%; Pred. No. 1.7e-62;
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY 96 QLGISFSNLSRITSGTTIIILASTTPGVKSNLQPTVTKTKNTTTOTOPSPKPTTKORONK 155
DB 490 QLGISFSNLSRITSGTTIIILASTTPGVKSNLQPTVTKTKNTTTOTOPSPKPTTKORONK 549
QY 156 PNKPNNDFHFEVNFVPCSCSNPTCWAICRIPNKKPGKKTTPKPTKPTTKKDL 215
DB 550 PSKPNNDPHEFVNFVPCSCSNPTCWAICRIPNKKPGKKTTPKPTKPTTKKDL 609
QY 216 KPQTTKKEVPTTKPEEPTINTTKNTITTLTNNTGNPKLTSQMEFHTSSSEGNLS 275
DB 610 KPQTTKKEVPTTKPEEPTINTTKNTITTLTNNTGNPKLTSQMEFHTSSSEGNLS 669
QY 276 PSQVSTTSE 284
DB 670 PSQVNISSQ 678
XX
RESULT 9
AAB68336
ID AAB68336 standard; protein: 299 AA.
XX
AC AAB68336;
XX
DT 09-JUL-2001 (first entry)
XX
DE Amino acid sequence of RSV G-protein.
XX
KW Respiratory syncytial virus; RSV; G-protein; annexin II; L-selectin;
XX RSV infection.
XX
OS Respiratory syncytial virus.
XX
XX WO200129054-A2.
XX
XX 26-APR-2001.
XX
XX 23-OCT-2000; 2000WO-GB04084.
XX
XX 21-OCT-1999; 99GB-0024990.
XX

```



PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;  
 PI Uhlen M;  
 XX  
 DR WPI: 1996-251766/25.  
 DR N-PSDB: AAT31647.  
 XX  
 PT Enhancing immunogenicity by coupling immunogen to serum  
 PT albumin-binding protein - useful for preparing improved vaccines,  
 PT e.g. against Respiratory Syncytial Virus  
 XX  
 PS Disclosure: Page 81-83; 102pp; French.  
 XX  
 CC The present sequence comprises a 219 amino acid fragment of  
 CC Streptococcal protein G, designated BB, fused to an immunogen,  
 CC designated G2A, derived from amino acids 130-230 of protein G from  
 CC Respiratory Syncytial Virus sub-group A or B. The BB fragment is  
 CC able to bind to human serum albumin and thereby enhance immunogenicity  
 CC of any antigen, hapten or immunogen that is covalently coupled to it.  
 CC In this specific example, the BB fragment was found to induce T helper  
 CC memory cells, leading to production of anti-G2A antibodies by stimulated  
 CC B cells.  
 XX  
 SO Sequence 349 AA:  
 Query Match 36.4%; Score 562; DB 17; Length 349;  
 Best Local Similarity 78.1%; Pred. No. 3.3e-36;  
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;  
 QY 103 NLSSETTSQTT--IIASTT-----PGVKSNILOPTTVTKNTTTTQTOTPSKPTTKORONKP 155  
 DB 216 NGKTLGSETTTEAVDAATARSFNFPILNSM--TVTKTKNTTTTQTOTPSKPTTKORONKP 272  
 QY 156 PNKPNNDHFEEVFNFPVCSICSNPTCAICRIIPNKKPGKTTTKPTKPTFKTKKDL 215  
 DB 273 PNKPNNDHFEEVFNFPVCSICSNPTCAICRIIPNKKPGKTTTKPTKPTFKTKKDL 332  
 QY 216 KPQTTRKREVPPTTKPTE 232  
 DB 333 KPQTTRKREVPPTTKPVD 349

RESULT 12  
 AAB68028  
 ID AAB68028 standard; Protein; 349 AA.  
 XX  
 AC AAB68028;  
 XX  
 DT 29-JUN-2001 (first entry)  
 XX  
 DE Amino acid sequence of fusion protein comprising 2 G protein fragments.  
 XX  
 KW Alphabetic ammonium salt; immunogen; antigen; syncytial virus infection;  
 KW G protein; fusion protein.  
 XX  
 OS Synthetic.  
 OS Streptococcus sp.  
 OS Respiratory syncytial virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 120..230  
 FT /note="G protein fragment of VRS"  
 XX  
 PN FR2798292-A1.  
 XX  
 PD 16-MAR-2001.  
 XX  
 PF 09-SEP-1999; 99FR-0011284.  
 XX  
 PR 09-SEP-1999; 99FR-0011284.  
 XX  
 PA (FABR ) FABRE MEDICAMENT SA PIERRE.

XX  
 PI Beck A, Klingner C, Nguyen TN;  
 XX  
 DR WPI: 2001-267782/28.  
 DR N-PSDB: AAF84711.  
 XX  
 PT Use of quaternary aliphatic ammonium salt and immunogen or antigen to  
 PT combat respiratory syncytial virus infections  
 XX  
 PS Claim 16; Page 22-23; 35pp; French.  
 XX  
 CC The specification describes a method which uses quaternary aliphatic  
 CC ammonium salts together with an immunogen or antigen to treat syncytial  
 CC virus infections. The combination of the salt with the antigen or  
 CC immunogen improves immunogenicity and equilibrates the Th1/Th2 immune  
 CC response. The method is used for the treatment of respiratory syncytial  
 CC virus infections. The present sequence represents a fusion protein,  
 CC comprising an albumin binding domain of the G protein of Streptococcus  
 CC sp. fused to a G protein fragment of respiratory syncytial virus (VRS,  
 CC long version). The fusion protein is used as an antigen in the method  
 CC of the invention.  
 XX  
 SO Sequence 349 AA:  
 Query Match 36.4%; Score 562; DB 22; Length 349;  
 Best Local Similarity 78.1%; Pred. No. 3.3e-36;  
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;  
 QY 103 NLSSETTSQTT--IIASTT-----PGVKSNILOPTTVTKNTTTTQTOTPSKPTTKORONKP 155  
 DB 216 NGKTLGSETTTEAVDAATARSFNFPILNSM--TVTKTKNTTTTQTOTPSKPTTKORONKP 272  
 QY 156 PNKPNNDHFEEVFNFPVCSICSNPTCAICRIIPNKKPGKTTTKPTKPTFKTKKDL 215  
 DB 273 PNKPNNDHFEEVFNFPVCSICSNPTCAICRIIPNKKPGKTTTKPTKPTFKTKKDL 332  
 QY 216 KPQTTRKREVPPTTKPTE 232  
 DB 333 KPQTTRKREVPPTTKPVD 349

RESULT 13  
 AAR8253  
 ID AAR8253 standard; peptide; 101 AA.  
 XX  
 AC AAR8253;  
 XX  
 DT 15-MAY-1996 (first entry)  
 XX  
 DE RSV subgroup A wild type residues 130-230 (G2A clone).  
 XX  
 KW Immunogenic peptide; protein G; respiratory syncytial virus; adjuvant;  
 KW carrier protein; membrane lipopolysaccharide; LPS; Klebsiella pneumoniae;  
 KW divalent cation; detergent; anion-exchange chromatography; infection.  
 XX  
 OS Respiratory syncytial virus.  
 OS  
 PN FR2718452-A1.  
 XX  
 PD 13-OCT-1995.  
 XX  
 PF 06-APR-1994; 94FR-0004009.  
 XX  
 PR 06-APR-1994; 94FR-0004009.  
 XX  
 PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Bausant T, Binz H, Thien NN, Trudel M;  
 XX  
 DR WPI: 1995-353189/46.  
 DR N-PSDB: AAT03486.  
 XX  
 PT New respiratory syncytial virus polypeptide(s) for vaccine prodn.

PT esp. by conjugation with new *Klebsiella pneumoniae* p40 protein  
XX  
XX Disclosure: Page 23; 38pp; French.  
PS  
XX The sequence of the wild type amino acid residues 130-230 from the  
CC protein G of the respiratory syncytial virus (RSV) subgroups A. The  
CC sequence was used to synthesise the immunogenic peptides AAR8245-52  
CC based on residues 174-187 of the subgroup A and B RSVs. The peptides are  
CC pref. conjugated to a novel carrier protein (the p40 protein; see  
CC AAR8237) derived from a membrane lipopolysaccharide (LPS) fraction from  
CC *Klebsiella pneumoniae*. The LPS fraction was isolated by precipitating the  
CC *Klebsiella* membrane LPSs with a divalent cation and detergents,  
CC subjecting the recovered proteins to anion-exchange chromatography to  
CC obtain an immunological adjuvant and linking the p40 protein to the above  
CC peptides. The conjugates are useful in the treatment of RSV A or B  
CC infections.  
XX  
SQ Sequence 101 AA;  
Query Match 36.1%; Score 558; DB 16; Length 101;  
Best Local Similarity 99.0%; Pred. No. 1.5e-36;  
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 130 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVFNFPSCISCSNPTCAICKRI 189  
DB 1 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVFNFPSCISCSNPTCAICKRI 60  
OY 190 PNKKPGKTTTKPTKPKPTFKTKDKLPQTTRKREVPPTTKP 230  
DB 61 PNKKPGKTTTKPTKPKPTFKTKDKHKPQTTRKREVPPTTKP 101  
RESULT 14  
AAR95610  
ID AAR95610 standard; peptide; 101 AA.  
XX  
XX AAR95610;  
AC  
XX  
XX 11-FEB-1997 (first entry)  
DT  
XX  
XX RSV subgp. A protein G wild type amino acids 130-230.  
DE  
XX  
XX Respiratory syncytial virus; protein G; heterologous protein; infection;  
KW cell surface; *Staphylococcus xylosum*; *Staphylococcus carnosus*; vaccine;  
KW expression vector; shuttle vector; hydrophobic region; immunogenic;  
KW streptococcal protein G serum albumin binding domain; conformation.  
XX  
XX Synthetic.  
OS  
XX  
XX WO9614409-A1.  
XX  
XX 17-MAY-1996.  
PD  
XX  
XX 07-NOV-1995; 95WO-FR01464.  
PF  
XX  
XX 07-NOV-1994; 94FR-0013307.  
PR  
XX  
XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
PA  
XX  
XX Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;  
PI  
XX  
XX WPI; 1996-251759/25.  
DR  
XX  
XX N-PSDB; AAT27066.  
DR  
XX  
XX Secretion of recombinant hydrophobic peptide analogue - esp. for  
PT secretion of respiratory syncytial virus antigenic peptide(s)  
PT suitable for use in oral vaccines  
PT  
XX  
XX Claim 12; Page 20; 47pp; French.  
PS  
XX This is the sequence of the wild type protein G amino acid 130-230 from  
CC the respiratory syncytial virus (RSV) subgroup A. The sequence is used  
CC in a method for producing heterologous proteins on the surface of a cell

CC pref. a non-pathogenic micro-organism such as *Staphylococcus xylosum* or  
CC *S. carnosus*. The coding sequence was generated synthetically by  
CC annealing and ligating the oligonucleotides AAT4252-74. The complete  
CC ligated sequence was inserted into the expression shuttle vector pSF/BRXM  
CC such that the 62 sequence (encoding amino acids 130-230) is fused to the  
CC sequence encoding the streptococcal protein G serum albumin binding  
CC domain (BB). The encoded peptides may have modifications in their  
CC hydrophobic regions which do not affect their activity but allow them to  
CC traverse the cell membrane and be displayed on the cell surface in the  
CC correct immunogenic conformations. Cells carrying such peptides can be  
CC used as vaccines against RSV infections.  
XX  
SQ Sequence 101 AA;  
Query Match 36.1%; Score 558; DB 17; Length 101;  
Best Local Similarity 99.0%; Pred. No. 1.5e-36;  
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 130 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVFNFPSCISCSNPTCAICKRI 189  
DB 1 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVFNFPSCISCSNPTCAICKRI 60  
OY 190 PNKKPGKTTTKPTKPKPTFKTKDKLPQTTRKREVPPTTKP 230  
DB 61 PNKKPGKTTTKPTKPKPTFKTKDKHKPQTTRKREVPPTTKP 101  
RESULT 15  
AAR95616  
ID AAR95616 standard; Protein; 101 AA.  
XX  
XX AAR95616;  
AC  
XX  
XX 10-FEB-1997 (first entry)  
DT  
XX  
XX RSV sub-group A wild type protein G residues 130-230.  
DE  
XX  
XX Heterologous protein; cell surface; *Staphylococcus xylosum*; protein G;  
KW *Staphylococcus carnosus*; respiratory syncytial virus; wild type; mutant;  
KW expression plasmid; fusion protein; streptococcal.  
XX  
XX Human respiratory syncytial virus.  
OS  
XX  
XX WO9614418-A1.  
XX  
XX 17-MAY-1996.  
PD  
XX  
XX 07-NOV-1995; 95WO-FR01465.  
PF  
XX  
XX 07-NOV-1994; 94FR-0013309.  
PR  
XX  
XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
PA  
XX  
XX Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;  
PI  
XX  
XX WPI; 1996-251768/25.  
DR  
XX  
XX N-PSDB; AAT27073.  
DR  
XX  
XX Expression of respiratory syncytial virus protein G at cell surface  
PT - of bacteria non-pathogenic for mammals, useful in orally active  
PT vaccines  
PT  
XX  
XX Claim 12; Page 13-14; 38pp; French.  
PS  
XX The invention relates to the prodn. of a heterologous protein on the  
CC surface of a cell, pref. *Staphylococcus xylosum* or *S. carnosus*, contg. a  
CC fragment of the respiratory syncytial virus comprising residues 130-230  
CC of the RSV protein G. This sequence is through to be the wild type  
CC residues 130-230 from the RSV subgp. A. The heterologous protein may  
CC also contain the subgp. B sequence (AAT27074). These sequences may be  
CC mutated to replace the Cys residues at pos. 173 and 186 by residues  
CC unable to form a disulphide bridge esp. Ser residues (see AAT27075-6).  
CC Other mutations include substitution of the Phe residues at pos. 163,

CC 165, 168 and/or 170 by Ser residues. The sequences encoding the wild  
 CC type or mutant RSV residues 130-230 are inserted into plasmid  
 CC PSE/mp18BBXM to create plasmid PSE/G2BBXM (wild type sequence) or  
 CC PSE/G2subBBXM (mutant sequence). The sequence is produced from PSE/BBXM  
 CC as a fusion protein where the fusion is with the streptococcal protein G  
 CC serum albumin binding domain. These plasmids are introduced into  
 CC S. xyloosus or S. carnosus for expression of the wild type or mutant  
 CC protein.  
 CC  
 XX

SO Sequence 101 AA;

Query Match 36.1%; Score 558; DB 17; Length 101;  
 Best Local Similarity 99.0%; Pred. No. 1.5e-36;  
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 130 TVKTKNTTQTOTPSKPTTKORONKPPNKNDFHFEVNFVPCISCSNPTGWAICKRI 189  
 |||||||  
 Db 1 TVKTKNTTQTOTPSKPTTKORONKPPNKNDFHFEVNFVPCISCSNPTGWAICKRI 60  
 OY 190 PNKKPGKTTTTPKPKTKTKKDLKPOTTKPREVPTTKP 230  
 |||||||  
 Db 61 PNKKPGKTTTTPKPKTKTKKDLKPOTTKPREVPTTKP 101

Search completed: May 1, 2003, 13:44:22  
 Job time : 38.6717 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:43:28 ; Search time 15.7434 Seconds  
(Without alignments)  
556.934 Million cell updates/sec

Title: US-09-462-816-2  
Perfect score: 1544  
Sequence: 1 MSKMKDQRTAKTEKTDLT.....VSTSPHPSPSPSPVTRQ 298

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544	100.0	298	2	US-08-467-963C-8
2	1544	100.0	298	2	US-08-838-189D-8
3	1544	100.0	298	3	US-08-852-344D-8
4	1544	100.0	298	3	US-08-344-639E-8
5	1544	100.0	298	4	US-08-467-969A-8
6	1544	100.0	298	4	US-08-467-961A-8
7	1544	100.0	298	4	US-08-001-554A-8
8	901	58.4	681	6	5194595-19
9	298.5	19.1	263	5	PCT-US91-08177-13
10	222	14.4	37	3	US-08-793-792-12
11	193	12.5	32	3	US-08-793-792-8
12	188	12.2	216	3	US-08-928-361B-8
13	188	12.2	1837	3	US-08-928-361B-5
14	187.5	12.1	1721	3	US-08-700-651-5
15	187.5	12.1	1721	3	US-08-928-361B-6
16	183.5	11.9	216	3	US-08-928-361B-27
17	174	11.3	28	3	US-08-793-792-4
18	173.5	11.2	249	3	US-08-700-651-15
19	173.5	11.2	249	3	US-08-928-361A-20
20	169	10.9	36	4	US-09-082-279B-871
21	169	10.9	36	4	US-09-315-304B-871
22	164.5	10.7	2476	2	US-08-276-967-2
23	162	10.5	941	4	US-07-757-022B-14
24	162	10.5	1022	4	US-07-757-022B-84
25	162	10.5	1038	4	US-07-757-022B-74
26	162	10.5	1049	4	US-07-757-022B-58
27	162	10.5	1140	4	US-07-757-022B-104

28	162	10.5	1270	4	US-07-757-022B-44	Sequence 44, Appl
29	162	10.5	1311	4	US-07-757-022B-42	Sequence 42, Appl
30	162	10.5	1313	4	US-07-757-022B-142	Sequence 142, App
31	162	10.5	1314	4	US-07-757-022B-50	Sequence 50, Appl
32	162	10.5	1320	4	US-07-757-022B-46	Sequence 46, Appl
33	162	10.5	1320	4	US-07-757-022B-60	Sequence 60, Appl
34	162	10.5	1354	4	US-07-757-022B-48	Sequence 48, Appl
35	162	10.5	1361	4	US-07-757-022B-40	Sequence 40, Appl
36	162	10.5	1363	4	PCT-US95-04611A-19	Sequence 52, Appl
37	162	10.5	1404	4	US-07-757-022B-52	Sequence 2, Appl
38	162	10.5	1404	4	US-07-757-022B-62	Sequence 62, Appl
39	161.5	10.5	878	4	US-09-556-706B-2	Sequence 2, Appl
40	159.5	10.3	907	3	US-08-783-774-2	Sequence 2, Appl
41	159.5	10.3	907	4	US-09-328-599A-1	Sequence 1, Appl
42	159.5	10.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl
43	158	10.2	175	3	US-08-700-651-12	Sequence 12, Appl
44	158	10.2	175	3	US-08-928-361B-17	Sequence 17, Appl
45	151	9.8	37	3	US-08-793-792-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-08-467-963C-8  
; Sequence 8, Application US/08467963C  
; Patent No. 5968776  
GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH COMBINES PROTECTION AGAINST  
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Slim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,963C  
; FILING DATE:  
CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/838,189  
; FILING DATE: 16-APR-1997  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-467-963C-8



TOPOLOGY: linear  
US-08-852-344D-8  
Query Match 100.0%; Score 1544; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.7e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNLSVAQITLSILAMISTSLITTAII 60  
Db 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNLSVAQITLSILAMISTSLITTAII 60  
Qy 61 FIASANHKVLTITTAIIODATSOIKNTTPYTLTOPDGLISFSLSLSTQTTTLLASTTP 120  
Db 61 FIASANHKVLTITTAIIODATSOIKNTTPYTLTOPDGLISFSLSLSTQTTTLLASTTP 120  
Qy 121 GVSNLQPTTVTKNTTTTOTOPOSKPTTKOROKPKNKPNNDHFVFNVPFCISGNP 180  
Db 121 GVSNLQPTTVTKNTTTTOTOPOSKPTTKOROKPKNKPNNDHFVFNVPFCISGNP 180  
Qy 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKDKLPQTTPKKEVPPTTKPTEEPPTINTTK 240  
Db 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKDKLPQTTPKKEVPPTTKPTEEPPTINTTK 240  
Qy 241 TINTTTLLTNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298  
Db 241 TINTTTLLTNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298  
RESULT 4  
US-08-344-639E-8  
Sequence 8, Application US/08344639E  
Patent No. 6033668  
GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/344,639E  
FILING DATE: 14-NOV-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-391 MIS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 298 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-344-639E-8  
Query Match 100.0%; Score 1544; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.7e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNLSVAQITLSILAMISTSLITTAII 60  
Db 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNLSVAQITLSILAMISTSLITTAII 60  
Qy 61 FIASANHKVLTITTAIIODATSOIKNTTPYTLTOPDGLISFSLSLSTQTTTLLASTTP 120  
Db 61 FIASANHKVLTITTAIIODATSOIKNTTPYTLTOPDGLISFSLSLSTQTTTLLASTTP 120  
Qy 121 GVSNLQPTTVTKNTTTTOTOPOSKPTTKOROKPKNKPNNDHFVFNVPFCISGNP 180  
Db 121 GVSNLQPTTVTKNTTTTOTOPOSKPTTKOROKPKNKPNNDHFVFNVPFCISGNP 180  
Qy 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKDKLPQTTPKKEVPPTTKPTEEPPTINTTK 240  
Db 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKDKLPQTTPKKEVPPTTKPTEEPPTINTTK 240  
Qy 241 TINTTTLLTNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298  
Db 241 TINTTTLLTNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298  
RESULT 5  
US-08-467-969A-8  
Sequence 8, Application US/08467969A  
Patent No. 6168786  
GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
TITLE OF INVENTION: CHIMERIC IMMUNOGENS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,969A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-475 MIS:dh  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 298 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-467-969A-8

Query Match 100.0%; Score 1544; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.7e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKNQDRTAKTLEKTDLTNLHLFISSGLYKLNKLSVAQITLSILAMISTSLITTAI 60  
|||||  
DB 1 MSKKNQDRTAKTLEKTDLTNLHLFISSGLYKLNKLSVAQITLSILAMISTSLITTAI 60  
QY 61 FIASANHKVLTFTAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTTTILASTP 120  
|||||  
DB 61 FIASANHKVLTFTAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTTTILASTP 120  
QY 121 GVSANLQPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNMP 180  
|||||  
DB 121 GVSANLQPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNMP 180  
QY 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKPEVPTTKPEEPTINTTK 240  
|||||  
DB 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKPEVPTTKPEEPTINTTK 240  
QY 241 TTTTTLTNTTGNPKLTISQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298  
|||||  
DB 241 TTTTTLTNTTGNPKLTISQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298

## RESULT 6

US-08-467-961A-8  
Sequence 8, Application US/08467961A  
Patent No. 6171783

GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: Ewasysbyn, Mary E  
TITLE OF INVENTION: Chimeric Immunogens  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6TH Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,961A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-476 Mib:hb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 298 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-467-961A-8

Query Match 100.0%; Score 1544; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.7e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKNQDRTAKTLEKTDLTNLHLFISSGLYKLNKLSVAQITLSILAMISTSLITTAI 60  
|||||  
DB 1 MSKKNQDRTAKTLEKTDLTNLHLFISSGLYKLNKLSVAQITLSILAMISTSLITTAI 60  
QY 61 FIASANHKVLTFTAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTTTILASTP 120  
|||||  
DB 61 FIASANHKVLTFTAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTTTILASTP 120  
QY 121 GVSANLQPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNMP 180  
|||||  
DB 121 GVSANLQPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNMP 180  
QY 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKPEVPTTKPEEPTINTTK 240  
|||||  
DB 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKPEVPTTKPEEPTINTTK 240  
QY 241 TTTTTLTNTTGNPKLTISQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298  
|||||  
DB 241 TTTTTLTNTTGNPKLTISQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298

## RESULT 7

US-08-001-554A-8

Sequence 8, Application US/08001554A  
Patent No. 6225091

GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: Ewasysbyn, Mary E  
TITLE OF INVENTION: Chimeric Immunogens  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/001,554A  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-286  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS

; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
US-08-001-554A-8

Query Match 100.0%; Score 1544; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1,7e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSIAMIISTLIITAI 60  
DB 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSIAMIISTLIITAI 60  
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTTP 120  
DB 61 FIASANKHVTLLTAIIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTTP 120  
QY 121 GKSNIQPTTVKTKNTTQTOPSKPTTKORONKPPNKNNDFEVEFVPCISCSNP 180  
DB 121 GKSNIQPTTVKTKNTTQTOPSKPTTKORONKPPNKNNDFEVEFVPCISCSNP 180  
QY 181 TCWAICKRIPNKKPGKKTTPKPKPTTKKDLKPOTTKREVPPTTKPTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKKTTPKPKPTTKKDLKPOTTKREVPPTTKPTEPTINTTK 240  
QY 241 TTTTTLTLLNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSSPPTTRQ 298  
DB 241 TTTTTLTLLNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSSPPTTRQ 298

RESULT 8  
5194595-19  
; Patent No. 5194595  
; APPLICANT: WATHEN, MICHAEL W.  
; TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING  
; IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY  
; SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/543,780  
; FILING DATE: 31-OCT-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 137,387  
; FILING DATE: 23-DEC-1987  
; SEQ ID NO: 19:  
; LENGTH: 681  
5194595-19

Query Match 58.4%; Score 901; DB 6; Length 681;  
Best Local Similarity 91.0%; Pred. No. 1.3e-73;  
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 96 QLGISFNSLSEITSTIILASTTPGVKSNLQPTTVKTKNTTQTOPSKPTTKORONK 155  
DB 490 QLGISFNSLSEITSTIILASTTPGVKSNLQPTTVKTKNTTQTOPSKPTTKORONK 155  
QY 156 PNKPNNDHFVFNFPVPCISNNPCTCAICKRIPNKKPGKKTTPKPKPTTKKDL 215  
DB 550 PSKRNDHFVFNFPVPCISNNPCTCAICKRIPNKKPGKKTTPKPKPTTKKDL 215  
QY 216 KPOTTKREVPPTTKPTEPTINTTKNTITLTLNNTGNKRLTSOMETHSTSSSEGNLS 275  
DB 610 KPOTTKREVPPTTKPTEPTINTTKNTITLTLNNTGNKRLTSOMETHSTSSSEGNLS 275  
QY 276 PSQVSTSE 284  
DB 670 PSQVSTSE 284

RESULT 9  
PCT-US91-08177-13  
; Sequence 13, Application PC/TUS9108177  
; GENERAL INFORMATION:  
; APPLICANT: Samal, Siba K  
; TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
; STREET: 1201 New York Avenue N.W., suite 1000  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/08177  
; FILING DATE: 19911104  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/608,937  
; FILING DATE: 05-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hignett, David W  
; REGISTRATION NUMBER: 30,265  
; REFERENCE/DOCKET NUMBER: 20509-96711  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-962-4854  
; TELEFAX: 202-962-8300

INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 263 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US91-08177-13

Query Match 19.1%; Score 295.5; DB 5; Length 263;  
Best Local Similarity 31.7%; Pred. No. 3.2e-19;  
Matches 86; Conservative 35; Mismatches 123; Indels 27; Gaps 7;

QY 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSIAMIISTLIITAI 60  
DB 1 MSNHTHPKFKRTYLKRAKASKRYFLVGLSCLYKFNLSLVQVLTALSLAMITLTVLITAI 60  
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTTP 120  
DB 61 YISVGNKAKAPTSKPTTQOQOPONHTPLPTEHNKSTHTSTQSTLSPPMI--DTIS 118  
QY 121 GVK-----SNLQPTTVKTKNT--TTOTOPSKPTTKORONKPPNKPNNDHFVFNFPVCS 174  
DB 119 GTTYGHPINRTONKIKISQSTPLATKRLPINP-----ESNPEHNDHNNSTQLPHVPCS 174  
QY 175 ICSNPNPCMAIC-----KRIPN-----KKPKKTTTPKPKPTTKKDLKPOTTTP 222  
DB 175 TCBSNPACSPLOIGLERASRAPPTTLKKAAPKPTTKPTKTTIYHRSPEAKLOTKN 234  
QY 223 KEVPTTKPTEPTINTTKNTITLTLNNT 253  
DB 235 TATPQOGLISSP---EHQTNOSTQISQHTS 262

RESULT 10  
US-08-793-792-12  
; Sequence 12, Application US/08793792  
; Patent No. 6077511  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Antigenic peptides derived from the











```
QY 122 ----VKSNIQPTVTKTKNTTT-----QTQPSKPTTKORONKPKPNNDH 164
      :      :      :      :      :      :      :      :      :
Db 1564 PPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      :      :      :      :      :      :      :      :      :
QY 165 FEVFNFPVPCISGNPPCMAICRIPNKKPGKTTTKPTKKPFKTKDKLPQTKRKE 224
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      :      :      :      :      :      :      :      :      :
QY 225 VPTTKPEEPTINTTKNTITTTLLTNNTGNPKLTSMETFSHSSSEGNLSPSOVSTSE 284
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTTPSSSTTTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 285 HPSQPSPPNTT 296
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 2
US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-922-217-1068

Query Match 12.5%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.7e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 71 LTTAIIODA--TSQIKNTTP-----TYLQDPOLGISFSNLSSEIT--SQTTTIIASTTPG 121
      :      :      :      :      :      :      :      :      :
Db 1504 MTPPTTPASTTTLPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      :      :      :      :      :      :      :      :      :
QY 122 ----VKSNIQPTVTKTKNTTT-----QTQPSKPTTKORONKPKPNNDH 164
      :      :      :      :      :      :      :      :      :
Db 1564 PPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      :      :      :      :      :      :      :      :      :
QY 165 FEVFNFPVPCISGNPPCMAICRIPNKKPGKTTTKPTKKPFKTKDKLPQTKRKE 224
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      :      :      :      :      :      :      :      :      :
QY 225 VPTTKPEEPTINTTKNTITTTLLTNNTGNPKLTSMETFSHSSSEGNLSPSOVSTSE 284
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTTPSSSTTTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 285 HPSQPSPPNTT 296
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 3
```

```
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-833-263-1068

Query Match 12.5%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.7e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 71 LTTAIIODA--TSQIKNTTP-----TYLQDPOLGISFSNLSSEIT--SQTTTIIASTTPG 121
      :      :      :      :      :      :      :      :      :
Db 1504 MTPPTTPASTTTLPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      :      :      :      :      :      :      :      :      :
QY 122 ----VKSNIQPTVTKTKNTTT-----QTQPSKPTTKORONKPKPNNDH 164
      :      :      :      :      :      :      :      :      :
Db 1564 PPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      :      :      :      :      :      :      :      :      :
QY 165 FEVFNFPVPCISGNPPCMAICRIPNKKPGKTTTKPTKKPFKTKDKLPQTKRKE 224
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      :      :      :      :      :      :      :      :      :
QY 225 VPTTKPEEPTINTTKNTITTTLLTNNTGNPKLTSMETFSHSSSEGNLSPSOVSTSE 284
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTTPSSSTTTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 285 HPSQPSPPNTT 296
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 4
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

Query Match
Best Local Similarity 11.0%; Score 169.5; DB 10; Length 1367;
Matches 62; Conservative 36; Mismatches 94; Indels 45; Gaps 8;

QY 84 KNTPTLYLQDDPOLGISFNLSITSTILLASTIPGV--KSNLOPTTVKKNIT--T 139
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 209 KSTSTTSSTSESTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 268
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 QTPSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKPKKT 199
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 KEKPTPTTSCKEKPTPPHD-----TPPCT-----KKKTTTCKTCTKKT 310
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 TKPTKPKPTTKKDLK-----QTKPKREVPTTKPTEE-----PINTTKTNTTTL 247
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 311 TTPVPPTSSSTTESSAPVPTPSSSTTESSAPVTSSSTTESSAPVPTPSSSTTESSAP 370
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 248 LTNNTGNKILNSQMETFHSTSEGNLSPQVSTSEHSQP-----SSPNTT 296
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 371 VTSST---ESSAPVTSSTTESSAPVPTPSSSTTESSAPVTSSTTESSAPVTS 424
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-216-393-341
; Sequence 341 Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 341
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393-341

Query Match
Best Local Similarity 10.5%; Score 162.5; DB 10; Length 288;
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;

QY 39 AQITSLIAMIISTLITAIIFIASANKVLTALIIDATSOI----- 83
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 AALSDAIDRVSDPLVSLDIVIREAAQAKFDLRLITDIASIGEGAMALMGEEAIFI 135
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 -----KNTPTLYLQDDPOLGISFNLSITSTILLASTIPGVKSNLOPTTVKTN 135
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 136 RPRRSKRGGKTTT-----TSSSTSTSTTTTSTTTTSTTTTSTTTTTP 181
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 TTTTQTSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKPG 195
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 TTTTSTTTTTPPT----- 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 KKTTPKPKPTTKKDLKPKQTKPKREVPTTKPTEEPINTTKTNTTLLTNNTGN 255
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 249
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 256 PKLTSQMETFHSTSEGNLSPQVSTSEHSQPSSPNTT 296
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 PTTT-----TTTEPTTTTSTTTTSTTTTSTTTTSTTTTSTT 284
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-216-393-344
; Sequence 344 Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393-344

Query Match
Best Local Similarity 10.5%; Score 162.5; DB 10; Length 288;
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;

QY 39 AQITSLIAMIISTLITAIIFIASANKVLTALIIDATSOI----- 83
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 AALSDAIDRVSDPLVSLDIVIREAAQAKFDLRLITDIASIGEGAMALMGEEAIFI 135
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 -----KNTPTLYLQDDPOLGISFNLSITSTILLASTIPGVKSNLOPTTVKTN 135
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 136 RPRRSKRGGKTTT-----TSSSTSTSTTTTSTTTTSTTTTSTTTTTP 181
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 TTTTQTSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKPG 195
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 TTTTSTTTTTPPT----- 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 KKTTPKPKPTTKKDLKPKQTKPKREVPTTKPTEEPINTTKTNTTLLTNNTGN 255
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 249
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 256 PKLTSQMETFHSTSEGNLSPQVSTSEHSQPSSPNTT 296
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 PTTT-----TTTEPTTTTSTTTTSTTTTSTTTTSTTTTSTT 284
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-10-124-557-14
; Sequence 14 Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
```



```
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 1038 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match 10.5%; Score 162; DB 12; Length 1038;
Best Local Similarity 24.6%; Pred. No. 0.00023;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDATSQIKNTTPYITLQDPOLGISFSNLSSEITTSQTTTILASTTPGVKSNLOPTTV 131
DB 190 TSKVLAKPTPKAEPTTK-----GPAITPKPEPTPTTKPEPASVTP--KEPTPTTI 237
QY 132 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSN 178
DB 238 KSAPTTPKEPAPTTTKSAPTTKPEPAPTTTKPEPAPTTKPEP-----APTITKEP 286
QY 179 NPTGMAICKRIPN-----KKPKKTTTKP-----TKKPTFKTTK 212
DB 287 APTTKSAPTTKPEPAPTTPKKPAPTTKPEPAPTTKPEPTTKPEPAPTTKPEPAPTTPK 346
QY 213 KDL-----KQPTTKPKPEVPTTKP-----TEEPTINTKINTNTTLLNNTGNPKLIS 260
DB 347 EPAPTPAKKPAPTTKPEPAPTTKPEPAPTTTKESPTTKPEPAPTTTKSAPTTTKPEAPT 406
QY 261 QMETFHTSSSEGNLSPSOVSTSEHPSQSSPPNNT 296
DB 407 TTKSAPTTKPE-----PSPTTKPEPAPTTTKPEPAPTT 438

RESULT 10
US-10-124-557-58
; Sequence 58, Application US/10124557
; Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
```

```
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 1049 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-124-557-58

Query Match 10.5%; Score 162; DB 12; Length 1049;
Best Local Similarity 24.6%; Pred. No. 0.00023;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDATSQIKNTTPYITLQDPOLGISFSNLSSEITTSQTTTILASTTPGVKSNLOPTTV 131
DB 233 TSKVLAKPTPKAEPTTK-----GPAITPKPEPTPTTKPEPASVTP--KEPTPTTI 280
QY 132 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSN 178
DB 281 KSAPTTPKEPAPTTTKSAPTTKPEPAPTTTKPEPAPTTKPEP-----APTITKEP 329
QY 179 NPTGMAICKRIPN-----KKPKKTTTKP-----TKKPTFKTTK 212
DB 330 APTTKSAPTTKPEPAPTTPKKPAPTTKPEPAPTTKPEPTTKPEPAPTTKPEPAPTTPK 389
QY 213 KDL-----KQPTTKPKPEVPTTKP-----TEEPTINTKINTNTTLLNNTGNPKLIS 260
DB 390 EPAPTPAKKPAPTTKPEPAPTTKPEPAPTTTKESPTTKPEPAPTTTKSAPTTTKPEAPT 449
QY 261 QMETFHTSSSEGNLSPSOVSTSEHPSQSSPPNNT 296
DB 450 TTKSAPTTKPE-----PSPTTKPEPAPTTTKPEPAPTT 481

RESULT 11
US-10-124-557-104
; Sequence 104, Application US/10124557
; Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
```



Db 407 TTKSAPTTPE---PSPTTKKEPAPTTKEPAPTT 438

RESULT 13

US-10-124-557-42  
Sequence 42, Application US/10124557  
Patent No. US20020137894A1  
GENERAL INFORMATION:

APPLICANT: Turner, Katherine

Clark, Stephen C.

Jacobs, Kenneth

Hewick, Rodney M.

Gesner, Thomas G.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Cserf, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 1311 amino acids

TYPE: amino acid

MOLECULE TYPE: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 42:

US-10-124-557-42

Query Match

Best Local Similarity 10.5%; Score 162; DB 12; Length 1311;

Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

Db 72 TTAIIOATSOIKNTPTTYLTQDPOLGISFSNLSSETTSOTTTILASTTPGVKSNLOPTTV 131

Db 231 TSKVLAKPTPKAETTK-----GPAITTPKEPTTPPKEPASTTP---KEPTPTTI 278

QY 132 K-----TKNTTTTQTOPSKPTTKORONKPPNNDHFVFNFPVCSISN 178

Db 279 KSAFTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APTTEKP 327

QY 179 NPTGMAICKRIPN-----KKPKCKTTTKP-----TKKPTTEKTK 212

Db 328 APTTTKAPTTTPKEPAPTTTKKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPK 387

QY 213 KDL-----KPTTKKEPAPTTK-----TEPTINTKNTITTTLLJNNTGKNPKLTS 260

Db 388 EPAPTPAKKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPKESPTTPKEPAPTTTKSAPTTTKAPPT 447

QY 261 OMETFSTSSSEGNLSPOVSSTTSEHPSQSSSPNTT 296

Db 448 TTKSAPTTPE---PSPTTKKEPAPTTKEPAPTT 479

RESULT 14

US-10-124-557-142  
Sequence 142, Application US/10124557  
Patent No. US20020137894A1  
GENERAL INFORMATION:

APPLICANT: Turner, Katherine

Clark, Stephen C.

Jacobs, Kenneth

Hewick, Rodney M.

Gesner, Thomas G.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Cserf, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 amino acids

TYPE: amino acid

MOLECULE TYPE: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 142:

US-10-124-557-142

Query Match

Best Local Similarity 10.5%; Score 162; DB 12; Length 1313;

Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

Db 72 TTAIIOATSOIKNTPTTYLTQDPOLGISFSNLSSETTSOTTTILASTTPGVKSNLOPTTV 131

Db 233 TSKVLAKPTPKAETTK-----GPAITTPKEPTTPPKEPASTTP---KEPTPTTI 280

QY 132 K-----TKNTTTTQTOPSKPTTKORONKPPNNDHFVFNFPVCSISN 178

Db 281 KSAFTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APTTEKP 329





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 1, 2003, 13:41:33 ; Search time 19.6792 Seconds

(without alignments)  
1455.750 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDDORTAKTLKMTDL.....VSTSEHPSPSPPTTRQ 298

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	99.1	298	1 MGNZRL	major surface glycoprotein
2	1418	91.8	298	1 MGNZ	major surface glycoprotein
3	1343	87.0	297	2 J01205	attachment protein
4	1343	87.0	297	2 J01208	attachment protein
5	1333	86.3	298	2 J05680	G protein - Human
6	1328	85.0	297	2 J01209	attachment protein
7	1324	85.8	297	2 J01206	attachment protein
8	1313	85.0	298	2 J01207	attachment protein
9	1307	84.7	297	2 J01204	attachment protein
10	767.5	49.7	292	1 MGNZ60	major surface glycoprotein
11	761	49.3	292	1 MGNZ18	major surface glycoprotein
12	425	27.5	467	1 VHNZ	nucleocapsid protein
13	299	19.4	250	2 P00768	glycoprotein G - b
14	299	19.4	257	1 MGNZBR	glycoprotein G - b
15	295.5	19.1	263	2 J02284	glycoprotein G - b
16	284	18.4	263	2 J02388	glycoprotein G - b
17	280	18.1	248	2 P00769	glycoprotein G - b
18	272	17.6	263	2 A48732	attachment glycoprotein
19	208	13.5	307	1 GSFF3	salivary glue protein
20	194	12.6	662	2 A45155	mucin FIM-C.1 - Af
21	193	12.5	3020	2 A43932	mucin 2 precursor, hypothetical
22	191	12.4	851	2 T22696	hypothetical protein
23	191	12.4	1161	2 S57180	probable membrane
24	190	12.3	322	2 A53715	apomucin precursor
25	188	12.2	1832	2 T31113	mucin-like glycoprotein
26	185.5	12.0	379	2 S50125	larval glue protein
27	183	11.9	217	2 S01358	salivary glue protein
28	183	11.9	354	2 T46740	microfilament shea
29	177	11.5	770	2 T22808	hypothetical protein

30	175	11.3	825	2 T29634	hypothetical protein
31	172.5	11.2	327	2 S20074	promastigote surface
32	172.5	11.2	1777	2 T34369	hypothetical protein
33	171.5	11.1	216	2 T51920	mucin - rhesus mac
34	170.5	11.0	279	2 S53363	mucin 5AC (clone J
35	170	11.0	3570	2 T45025	mucin MUC5B, trach
36	169.5	11.0	1367	1 S48478	glucan 1,4-alpha-g
37	168	10.9	402	2 E86185	hypothetical prote
38	167.5	10.8	294	2 A37232	mucin, tracheal (A
39	167.5	10.8	1630	2 A53577	ascites stialoglyco
40	164.5	10.7	2476	2 T34022	zonadhesin - pig
41	163	10.6	371	2 S20075	promastigote surfa
42	163	10.6	477	2 S53362	mucin 5AC (clone J
43	160.5	10.4	393	2 S62335	171-7 protein - fr
44	160.5	10.4	866	2 T45462	membrane glycoprot
45	160.5	10.4	867	2 T45463	membrane glycoprot

## ALIGNMENTS

## RESULT 1

MGNZRL major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: A32703, S12279

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: A32703; MUID:87289657; PMID:2441388

A:Accession: A32703

A:Molecule type: mRNA

A:Residues: 1-298 &lt;J0H&gt;

A:Cross-References: GB:M17212; NID:9333940; PIDN:AAA47411.1; PID:9333941

R:Garcia-Barreno, B.; Portela, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.

EMBO J. 9, 4181-4187, 1990

A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizat

A:Reference number: S12279; MUID:91065351; PMID:2249671

A:Accession: S12279

A:Molecule type: mRNA

A:Residues: 1-298 &lt;GAR&gt;

A:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted &lt;TMN&gt;

F:85,103,135,179,237,250,251,273,294/Binding site: carbohydrate (asn) (covalent) #sta

Query Match 99.1%; Score 1530; DB 1; Length 298

Best local similarity 99.0%; Pred. No. 2,4e-92;

Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	1	MSKNKDDORTAKTLKMTDLNLHLFISSGLYKLKLSVAQITISLAMIISTSLITAIT 60
DB	1	MSKNKDDORTAKTLKMTDLNLHLFISSGLYKLKLSVAQITISLAMIISTSLITAIT 60
QY	61	FIAANHKVLTITTAIIDATSOIKNTPTVLTODPQIGISFNSSETTSOTTTLASTP 120
DB	61	FIAANHKVLTITTAIIDATSOIKNTPTVLTODPQIGISFNSSETTSOTTTLASTP 120
QY	121	GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNMP 180
DB	121	GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNMP 180
QY	181	TCMAICRIPKPKKGGKTTTKPKPKPKTKKDKLQOTTTPKPKKVPPTTKTEPTINTTK 240
DB	181	TCMAICRIPKPKKGGKTTTKPKPKPKTKKDKLQOTTTPKPKKVPPTTKTEPTINTTK 240
QY	241	TNITTTLLTNNTGNPKLTSSOMETFSSEGNLSPSOVSTSPHPSPSPPTTRQ 298
DB	241	TNITTTLLTNNTGNPKLTSSOMETFSSEGNLSPSOVSTSPHPSPSPPTTRQ 298

## RESULT 2

MG22  
Major surface glycoprotein G - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 24-Sep-1999  
C:Accession: A94048; A93599; A04039  
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
A:Reference number: A94048; M01D:85216656; PMID:3858865  
A:Accession: A94048  
A:Molecule type: mRNA  
A:Residues: 1-298 <MER>  
A:Cross-references: GB:K01486; GB:K02719; GB:K03348; GB:K03349; GB:M1217; GH  
7.1; PID:333932  
A:Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
A:Note: this protein may carry 40-80 separate O-linked carbohydrate chains distributed a  
R:Satake, M.; Colligan, J.E.; Elango, N.; Morley, E.; Venkatesan, S.  
Nucleic Acids Res. 13, 7795-7812, 1985  
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
A:Reference number: A93599; M01D:86067198; PMID:4069997  
A:Accession: A93599  
A:Molecule type: mRNA  
A:Residues: 1-298 <SAT>  
A:Cross-references: GB:X03149; M1D:960997; PIDN:CA26928.1; PID:960998  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:85,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 1418; DB 1; Length 298;  
Best Local Similarity 93.3%; Pred. No. 4,4e-85;  
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
Db 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
OY 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
Db 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
OY 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
Db 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
OY 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298  
Db 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 3  
J01205  
Attachment protein - human respiratory syncytial virus (strain RSB1734)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: J01205  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: J01204; M01D:91374005; PMID:1895054  
A:Accession: J01205  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;  
Best Local Similarity 88.6%; Pred. No. 3.2e-80;  
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
Db 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
OY 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
Db 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
OY 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
Db 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
OY 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 296  
Db 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 296

RESULT 4  
J01208  
Attachment protein - human respiratory syncytial virus (strain RSB6256)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: J01208  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
A:Reference number: J01204; M01D:91374005; PMID:1895054  
A:Accession: J01208  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 87.0%; Score 1343; DB 2; Length 297;  
Best Local Similarity 88.5%; Pred. No. 3.2e-80;  
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
Db 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSTILLASTP 120  
OY 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
Db 121 GVSANLOPTVTKNTTQTOPSKPTTKORONKPPNNDFHEVFNPVPCISGNNP 180  
OY 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
Db 181 TCWAICRIIPNKKPKKTTTPKPKPTTKKDKLPQTTKPKREVPYTKPTEPTINTTK 240  
OY 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 296  
Db 241 TMTITLLTNNTGNPKLTISOMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 296

RESULT 5  
J05680

G protein - human respiratory syncytial virus  
C:Species: Human respiratory syncytial virus  
C:Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
C:Accession: J05680  
R:Gen: X.; Wang, Z.; Qian, Y.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
Chinese J. Virol. 12, 317-322, 1996  
A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain 15  
A:Reference number: J05680  
A:Accession: J05680  
A:Molecule type: mRNA  
A:Residues: 1-298 <GEN>  
A:Experimental source: strain B79  
A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
F:1-38/Domain: intracellular #status predicted <INT>  
F:39-66/Domain: transmembrane #status predicted <TM>  
F:67-298/Domain: extracellular #status predicted <EX>

Query Match 86.3%; Score 1333; DB 2; Length 298;  
Best Local Similarity 87.9%; Pred. No. 1.4e-79;  
Matches 262; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
Db 1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
QY 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
Db 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
QY 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
Db 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
QY 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 298  
241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 298  
Db 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 298

## RESULT 6

J01209  
attachment protein - human respiratory syncytial virus (strain RSB614)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: J01209  
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: J01204; MUID:91374005; PMID:1895054  
A:Accession: J01209  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 86.0%; Score 1328; DB 2; Length 297;  
Best Local Similarity 87.8%; Pred. No. 3e-79;

Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
Db 1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120

Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
QY 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
Db 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
QY 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
Db 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
QY 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296  
241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296  
Db 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296

## RESULT 7

J01206  
attachment protein - human respiratory syncytial virus (strain RSB5857)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: J01206  
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
A:Reference number: J01204; MUID:91374005; PMID:1895054  
A:Accession: J01206  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 85.8%; Score 1324; DB 2; Length 297;  
Best Local Similarity 87.8%; Pred. No. 5.5e-79;  
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
Db 1 MSKTKDORTAKTLEKTDNLHLFISSGLYKLNKSAOITLSILAMITSLITTAII 60  
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTIIASTTP 120  
QY 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
Db 121 GKSNLOPTTVKTKNTTQTQPSKPTTKORONKPPKNNDPFHEVNFVPCISCSNP 180  
QY 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
Db 181 TCMAICKRIPNKKPGKKTTPKTKPTKKTDLKPQTTPKKEVPTTKPTEPTINTTK 240  
QY 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296  
241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296  
Db 241 TTTTLLTNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSPSSPNTTQ 296

## RESULT 8

J01207  
attachment protein - human respiratory syncytial virus (strain RSB6190)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: J01207  
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
A:Reference number: J01204; MUID:91374005; PMID:1895054  
A:Accession: J01207  
A:Molecule type: mRNA

A:Residues: 1-298 <CAN>  
A:Note: the authors translated the codon ACC for residue 4 as Asn and AGC for residue 22  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-  
idren and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,250,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.0%; Score 1313; DB 2; Length 298;

Best Local Similarity 86.9%; Pred. No. 2.8e-78;  
Matches 259; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKTDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
QY 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
Db 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
QY 121 GVSXNLOPTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
Db 121 SVESIIQSTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
QY 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 240  
Db 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 240  
QY 241 TMTITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298  
Db 241 PMRTLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 9  
J01204  
attachment protein - human respiratory syncytial virus (strain NS642)

N:Alternate names: G protein

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: J01204

J:Gen. Virol. 72, 2091-2096, 1991

R:Can. P.A.: Matthews, D.A.; Pringle, C.R.

A:Title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: J01204; MUID:91374005; PMID:1895054

A:Accession: J01204

A:Molecule type: mRNA

A:Residues: 1-297 <CAN>

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-  
idren and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.7%; Score 1307; DB 2; Length 297;

Best Local Similarity 86.5%; Pred. No. 6.9e-78;  
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKNDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
QY 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
Db 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
QY 121 GVSXNLOPTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
Db 121 GVSXNLOPTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
QY 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 240  
Db 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 240  
QY 241 TMTITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298  
Db 241 TMTITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

QY 241 TMTITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 297  
Db 241 TMTITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 297

RESULT 10

MGN218  
major surface glycoprotein G - human respiratory syncytial virus (strain 8/60)

N:Alternate names: attachment glycoprotein G

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C:Accession: A37077

R:Sullender, W.M.; Anderson, K.; Wertz, G.W.

Virol. 178, 195-203, 1990

A:Title: The respiratory syncytial virus subgroup B attachment glycoprotein: analysis

A:Reference number: A37077; MUID:90357765; PMID:1697126

A:Accession: A37077

A:Molecule type: mRNA

A:Residues: 1-292 <SUL>

A:Cross-references: EMBL:M55633; NID:g333944; PIDN:AAA47413.1; PID:g333945

C:Genetics:

A:Gene: G

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:45-63/Domain: transmembrane #status predicted <TMN>

F:81,86,100,220,290/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.7%; Score 767.5; DB 1; Length 292;

Best Local Similarity 55.3%; Pred. No. 6.5e-43;  
Matches 162; Conservative 34; Mismatches 94; Indels 3; Gaps 3;

QY 1 MSKNDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
Db 1 MSKNDQRTAKTEKTWTLNHLFISSGLYKLNKSAQITSLIAMIISTSLITAI 60  
QY 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
Db 61 FIASANHKVTLTAAIIODATSOIKNTPTLYLQDPOLGISFNSLSEITSOQTTLIASTTP 120  
QY 121 GVSXNLOPTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
Db 121 GVSXNLOPTVYKTKNTTQTOPSKPTTKORONKPNNDHFEVFNVPSCISGNNP 180  
QY 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 239  
Db 181 TQWAIQKRIIPNKKPKKTKTKTKPKPKTKKDLKPKOTTKPKREVTPKPEEPTINTTK 239  
QY 240 KTNITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 292  
Db 240 KTNITLLTNTTGNPKLTSMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 292

RESULT 11

MGN218  
major surface glycoprotein G - human respiratory syncytial virus (strain 18537)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: B32703

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: B32703; MUID:87289657; PMID:2441388

A:Accession: B32703

A:Molecule type: mRNA

A:Residues: 1-292 <JOH>

A:Cross-references: GB:M17213; NID:g333942; PIDN:AAA47412.1; PID:g333943

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.3%; Score 761; DB 1; Length 292;

Best Local Similarity 54.9%; Pred. No. 1.7e-42;  
Matches 158; Conservative 35; Mismatches 93; Indels 2; Gaps 2;

QY 1 MSNKKDQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60  
1 MSKHKKQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60  
Db 1 MSKHKKQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60

QY 61 FIASANKVLTALIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTP 120  
1 FIASANKVLTALIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTP 120  
Db 61 FIASANKVLTALIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTP 120

QY 121 GYKSNLOPTVTKNTTQTOTOPSKPTTKQKRONKPPKNNDFHEVEVNFVPCISGNNP 180  
1 GYKSNLOPTVTKNTTQTOTOPSKPTTKQKRONKPPKNNDFHEVEVNFVPCISGNNP 180  
Db 121 GYKSNLOPTVTKNTTQTOTOPSKPTTKQKRONKPPKNNDFHEVEVNFVPCISGNNP 180

QY 181 TCWAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 239  
1 TCWAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 239  
Db 181 LCKSICKTIPSNKPKKPKPIKPTTKTKTKKDKPKTPAKMKKEIITNPAKKPKPLKTT 240

QY 240 KTNITTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPS 287  
1 KTNITTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPS 287  
Db 241 ERDTSISOSTVLDITTPKYTIQOOSLHSTPSSENPSTQIPTASE-PS 287

## RESULT 12

nuclEocapsid protein (version 2) - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999  
C:Accession: A04026  
R:Elango, N.; Venkatesan, S.  
Nucleic Acids Res. 11, 5941-5951, 1983  
A:Title: Amino acid sequence of respiratory syncytial virus capsid protein.  
A:Reference number: A04026; MUID:83299261; PMID:6310521  
A:Accession: A04026  
A:Molecule type: mRNA  
A:Residues: 1-467 <ELI>  
A:Cross-references: GB:X00001; MID:961215; PIDN:CAA24906.1; PID:961216  
C:Genetics:  
A:Gene: N  
C:Superfamily: respiratory syncytial virus nucleocapsid protein  
C:Keywords: nucleocapsid

Query Match 27.5%; Score 425; DB 1; Length 467;  
Best Local Similarity 90.2%; Pred. No. 1.7e-20;  
Matches 83; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 207 TFKTKKDLKPKQTPKPEPTTKPEPTNTTKNTTLLNNTGNPKLTSQMETFH 266  
1 TFKTKKDLKPKQTPKPEPTTKPEPTNTTKNTTLLNNTGNPKLTSQMETFH 266  
Db 376 TLKTKKDPKPKQTPKPEPTTKPEPTNTTKNTTLLNNTGNPKLTSQMETFH 435

QY 267 STSSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298  
1 STSSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298  
Db 436 STSSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 467

## RESULT 13

glycoprotein G - bovine respiratory syncytial virus (isolate FS-1) (fragment)  
C:Species: bovine respiratory syncytial virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999  
C:Accession: P00768  
R:Malipiedt, S.K.; Samal, S.K.  
J. Gen. Virol. 74, 2001-2004, 1993  
A:Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial virus.  
A:Reference number: J02284; MUID:93389461; PMID:8376974  
A:Accession: P00768  
A:Molecule type: mRNA  
A:Residues: 1-250 <MAL>  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:1-31/Domain: intracellular #status predicted <INT>  
F:32-59/Domain: transmembrane #status predicted <TM>

F:60-250/Domain: extracellular #status predicted <EXT>  
F:78,156,226,241,244/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 2; Length 250;  
Best Local Similarity 30.3%; Pred. No. 1.3e-12;  
Matches 91; Conservative 35; Mismatches 104; Indels 70; Gaps 9;

QY 11 KTEKWTWTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 70  
1 KTEKWTWTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 70  
Db 4 KTKKAMKASRYFVIGSLTKPKFKSLVQTAISTLMTITLSLITTAII 63

QY 71 LTTAIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTPGVKSNLOPTT 130  
1 LTTAIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTPGVKSNLOPTT 130  
Db 64 PTKPTIQOQTPQNNHSPFTENY-----KTHHSIOSTTISQIINDITTR 111

QY 131 VKTKNTTQTOTOPSK-----PTTKQ-----KONKPPKNNDFHEVEVNFVPCISGNNP 178  
1 VKTKNTTQTOTOPSK-----PTTKQ-----KONKPPKNNDFHEVEVNFVPCISGNNP 178  
Db 112 GTTGHSTDETOSSKRKISQSTLPTTRKPPINPSESNNPDHNNQSTLPPVPCSTCEG 171

QY 179 NPTCAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 238  
1 NPTCAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 238  
Db 172 NLACLSIQVGPGR-----APSRAPT-TLKTKTKPKTK-----KP----- 207

QY 239 KTNITTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298  
1 KTNITTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298  
Db 208 -----IKTTI-----HHRTSPEAKLQPKN-NPAAPQGLISSPENHTNQ 245

## RESULT 14

major surface glycoprotein G - bovine respiratory syncytial virus (strain 391-2)  
N:Alternate names: attachment glycoprotein G  
C:Species: bovine respiratory syncytial virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: A36408  
R:Leich, R.A.; Anderson, K.; Wertz, G.W.  
J. Virol. 64, 5559-5569, 1990  
A:Title: Nucleotide sequence analysis and expression from recombinant vectors demonstr  
y syncytial virus.  
A:Reference number: A36408; MUID:91012801; PMID:2214024  
A:Accession: A36408  
A:Molecule type: mRNA  
A:Residues: 1-257 <LE>  
A:Cross-references: GB:M58307; MID:9210830; PIDN:AAA42810.1; PID:9210831  
C:Genetics:  
A:Gene: G  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:45-62/Domain: transmembrane #status predicted <TM>  
F:3,85,127,149,233,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 1; Length 257;  
Best Local Similarity 31.6%; Pred. No. 1.4e-12;  
Matches 92; Conservative 33; Mismatches 120; Indels 46; Gaps 8;

QY 1 MSNKKDQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60  
1 MSNKKDQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60  
Db 1 MSNKKDQRTAKTEKTWDTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLITTAII 60

QY 61 FIASANKVLTALIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTP 120  
1 FIASANKVLTALIIODATSQIKNTPTLYLTODPOLGISFSNLSSETISQTTIIASTTP 120  
Db 61 YISVGNKAKRPTSKPTIQOQPNHNSPFTENY-----KTHHSIOSTTISQIINDITTR 108

QY 121 GYKSNLOPTVTKNTTQTOTOPSK-----PTTKQ-----KONKPPKNNDFHEVEVNFVPCISGNNP 178  
1 GYKSNLOPTVTKNTTQTOTOPSK-----PTTKQ-----KONKPPKNNDFHEVEVNFVPCISGNNP 178  
Db 109 SOLINTDTTGTGTYGSHSTNETQNKIKGOSTLPATKPPINPSESGIPPEHNDHNNFQTL 168

QY 169 NPYPCISGNNPPTCAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 228  
1 NPYPCISGNNPPTCAICKRIPNKKPGKKTTPKPKPTTKT-KKDLKPGQTKPKKVPPTKPEEPTINT 228  
Db 169 PYVPCSTCEGNLACLSI-----IETRAPSRAPT-TLKTKTKPKTK----- 212

QY 229 KPEEPTINTTKNTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOV 279  
1 KPEEPTINTTKNTTLLNNTGNPKLTSQMETFHSSTSEGNLSPSOV 279

Db 213 KPTK--TTIHRTSPETKLPKNNATP---QQGILSTEHNTOSTOI 257

## RESULT 15

JQ2284

glycoprotein G - bovine respiratory syncytial virus (isolate A51908)

C:Species: bovine respiratory syncytial virus

C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C:Accession: JQ2284

R:Mailpedd1, S.K.; Samal, S.K.

J. Gen. Virol. 74, 2001-2004, 1993

A:Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial

A:Reference number: JQ2284; MUID:93389461; PMID:8376974

A:Accession: JQ2284

A:Molecule type: mRNA

A:Residues: 1-263 &lt;MALD&gt;

A:Experimental source: isolate A51908

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:1-18/Domain: intracellular #status predicted &lt;CYT&gt;

F:39-66/Domain: transmembrane #status predicted &lt;TM&gt;

F:67-263/Domain: extracellular #status predicted &lt;EXT&gt;

F:127,163,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.1%; Score 295.5; DB 2; Length 263;

Best Local Similarity 31.7%; Pred. No. 2,3e-12;

Matches 86; Conservative 35; Mismatches 123; Indels 27; Gaps 7;

QY 1 MSKNKDORTAKTLEKTWDTLNHLFISSGLYKLNKSVAGITLSLAMIISTSLITAI 60

Db 1 MSNHTHHPRFKTLKRAMKASKYFIVGLSLCYKFNLSLVQATLSLAMITLTSIVITAI 60

QY 61 FIASANHVTLLTALIQDTSQIKNTPTPYLQDPQLGISFNSLSEITSTTILASTP 120

Db 61 YISVGNARAKPSTKPTTQTPQNNHTPLPTEHNNKSTHTSTOSTTLLSQPPNI--DTTS 118

QY 121 GVK-----SNLOPTVTKNT-TTOTOQSKPTKORONKPPNNDHFEVENFVPCS 174

Db 119 GTTYGHPINRQNRKIKSGSTPLATRKLPINL-----ESNPPENQDHNNSQTLPHVPCS 174

QY 175 IGSNNPTCAIC-----KRIPN-----KPKGKTTTKPTKPTFKTKDKLPQTKP 222

Db 175 TCEGNPACSPLOIGIERAPSRAPTITLKKAPKPKTKPTKTIYHRTSPKALQTKN 234

QY 223 KEVPTTKPEEPPTINTTKNITNTLTNNNT 253

Db 235 TAPPOGILSP---EHQTNOSTQISOHTS 262

Search completed: May 1, 2003, 13:47:16

Job time : 21.6792 secs

GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: May 1, 2003, 12:35:37 ; Search time 12.9321 Seconds  
(Without alignments)  
955.759 Million cell updates/sec

Title: US-09-462-816-2  
Perfect score: 1544  
Sequence: 1 MSKNKDQRTAKTEKTWDTL.....VSTSEHPQSSPPTTRQ 298

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	99.1	298	1	VGLG_HRSVL
2	1418	91.8	298	1	VGLG_HRSVA
3	1343	87.0	297	1	VGLG_HRSV3
4	1343	87.0	297	1	VGLG_HRSV6
5	1328	86.0	297	1	VGLG_HRSV7
6	1324	85.8	297	1	VGLG_HRSV4
7	1313	85.0	298	1	VGLG_HRSV5
8	1307	84.7	297	1	VGLG_HRSV2
9	767.5	49.7	292	1	VGLG_HRSV8
10	761	49.3	292	1	VGLG_HRSV1
11	309.5	20.0	263	1	VGLG_HRSV1
12	308.5	20.0	263	1	VGLG_HRSV4
13	299	19.4	257	1	VGLG_HRSV4
14	295.5	19.1	263	1	VGLG_HRSV2
15	294.5	19.1	263	1	VGLG_HRSV2
16	285	18.5	257	1	VGLG_HRSV5
17	284	18.4	263	1	VGLG_HRSV7
18	279	18.1	257	1	VGLG_HRSV7
19	270	17.5	257	1	VGLG_HRSV1
20	267	17.3	257	1	VGLG_HRSV4
21	208	13.5	307	1	SGS3_DROHE
22	198.5	12.9	2700	1	ZAN_HUMAN
23	194	12.6	662	1	MUC1_XENLA
24	193	12.5	5179	1	MUC2_HUMAN
25	191	12.4	1161	1	DAN4_YEAST
26	183	11.9	217	1	SGS3_DROST
27	171.5	11.1	2282	1	ZAN_RABIT
28	169.5	11.0	1367	1	AMYH_YEAST
29	164.5	10.7	2476	1	ZAN_PIG
30	160	10.4	263	1	SGS3_DROVA
31	159.5	10.3	907	1	VGP3_EBV
32	157.5	10.2	5376	1	ZAN_MOUSE
33	156.5	10.1	797	1	VGLX_HSVB

34	155	10.0	886	1	VGP3_EBVA8
35	155	10.0	1260	1	ALSI_CANAL
36	155	10.0	3178	1	YAG3_CAEFL
37	147	9.5	860	1	CH12_COCIM
38	146.5	9.5	467	1	P60_LISTIN
39	145	9.4	1075	1	FLO5_YEAST
40	144.5	9.4	1322	1	YAG3_YEAST
41	144	9.3	605	1	YHCB_YEAST
42	144	9.3	767	1	AMYH_SACDI
43	143.5	9.3	1251	1	YOUT3_CAEFL
44	143	9.3	484	1	P60_LISTMO
45	142.5	9.2	1758	1	YIR7_YEAST

## ALIGNMENTS

RESULT 1  
VGLG\_HRSVL STANDARD; PRT; 298 AA.  
AC P20895;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (subgroup A / strain long).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=67289657; Pubmed=2441388;  
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
RT "The G glycoprotein of human respiratory syncytial viruses of  
RT subgroups A and B: extensive sequence divergence between  
RT antigenically related proteins."  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M17212; AAA47411.1; -.  
DR PIR: A32703; MGNZRL.  
DR InterPro: IPR000925; Glycoprot.G.  
DR Pfam: PF00802; Glycoprotein\_G.1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37  
FT TRANSMEM 37  
FT POTENTIAL.  
FT CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 67 298  
FT CARBOHYD 103 298  
FT CARBOHYD 103 298  
FT CARBOHYD 135 135  
FT CARBOHYD 237 237  
FT CARBOHYD 250 250  
FT CARBOHYD 251 251  
FT CARBOHYD 294 294  
SQ SEQUENCE 298 AA; 32781 MW; B79FEFA4BA73B0E CRC64;  
Query Match 99.1%; Score 1530; DB 1; Length 298;  
Best Local Similarity 99.0%; Pred. No. 1,4e-93;  
Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSKNKDOPTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60  
DB 1 MSKNKDOPTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60  
QY 61 FIASANHKVLTPTALIODATSOIKNTPTLYTOPOLGISFNSLSEITSOITTLASTP 120  
DB 61 FIASANHKVLTPTALIODATSOIKNTPTLYTOPOLGISFNSLSEITSOITTLASTP 120  
QY 121 GVSXNLOPTVTKNTTOTOPTSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180  
DB 121 GVSXNLOPTVTKNTTOTOPTSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180  
QY 181 TCWAICRIPNKKPKGKTTTKPTKKPFTKTKDLKPOFTKPKREVPPTKPTKEPTINTTK 240  
DB 181 TCWAICRIPNKKPKGKTTTKPTKKPFTKTKDLKPOFTKPKREVPPTKPTKEPTINTTK 240  
QY 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSQSSPPTTRQ 298  
DB 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSQSSPPTTRQ 298  
RESULT 2  
VGLG\_HRSVA STANDARD: PRT: 298 AA.  
ID VGLG\_HRSVA STANDARD: PRT: 298 AA.  
AC P03423;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain A2).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11259;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=85216636; PubMed=3858865;  
RA Wert G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;  
RT "Nucleotide sequence of the G protein gene of human respiratory  
RT syncytial virus reveals an unusual type of viral membrane protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=86067198; PubMed=4069997;  
RA Satake M., Colligan J.E., Elango N., Norrby E., Venkatesan S.;  
RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel  
RT structure".  
RL Nucleic Acids Res. 13:7795-7812(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=95265253; PubMed=7747420;  
RA Connors M., Crowe J.E., Jr., Firestone C.Y., Murphy B.R., Collins P.L.;  
RT "A cold-passaged, attenuated strain of human respiratory syncytial  
RT virus contains mutations in the F and L genes".  
RL Virology 208:478-484(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=97187925; PubMed=9035372;  
RA Crowe J.E., Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,  
RA Murphy B.R.;  
RT "Acquisition of the ts phenotype by a chemically mutagenized cold-  
RT passaged human respiratory syncytial virus vaccine candidate results  
RT from the acquisition of a single mutation in the polymerase (L)  
RT gene".  
RL Virus Genes 13:269-273(1996).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M11486; AAB59857.1; -;  
DR EMBL: X03149; CAA26928.1; -;  
DR EMBL: U50362; AAB86663.1; -;  
DR EMBL: U50363; AAB86675.1; -;  
DR EMBL: U63644; AAC55969.1; -;  
DR PIR: A04039; MGNY.  
DR InterPro: IPR000925; Glycoprotein\_G.  
DR Pfam: PF00802; Glycoprotein\_G.1.  
KW Transmembrane; Glycoprotein.  
KW Cytoplasmic (POTENTIAL).  
FT DOMAIN 1 37  
FT TRANSMEM 38 66  
FT DOMAIN 67 298  
FT CARBOHYD 135 135  
FT CARBOHYD 237 237  
FT CARBOHYD 251 251  
SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;  
Query Match 91.8%; Score 1418; DB 1; Length 298;  
Best Local Similarity 93.3%; Pred. No. 3e-86;  
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
QY 1 MSKNKDOPTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60  
DB 1 MSKNKDOPTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60  
QY 61 FIASANHKVLTPTALIODATSOIKNTPTLYTOPOLGISFNSLSEITSOITTLASTP 120  
DB 61 FIASANHKVLTPTALIODATSOIKNTPTLYTOPOLGISFNSLSEITSOITTLASTP 120  
QY 121 GVSXNLOPTVTKNTTOTOPTSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180  
DB 121 GVSXNLOPTVTKNTTOTOPTSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180  
QY 181 TCWAICRIPNKKPKGKTTTKPTKKPFTKTKDLKPOFTKPKREVPPTKPTKEPTINTTK 240  
DB 181 TCWAICRIPNKKPKGKTTTKPTKKPFTKTKDLKPOFTKPKREVPPTKPTKEPTINTTK 240  
QY 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSQSSPPTTRQ 298  
DB 241 TMTITLLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSQSSPPTTRQ 298  
RESULT 3  
VGLG\_HRSV3 STANDARD: PRT: 297 AA.  
ID VGLG\_HRSV3 STANDARD: PRT: 297 AA.  
AC P27022;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb1734).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11253;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses".  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND



```

CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01205; J01205.
DR      InterPro: IPR000925; Glycoprot-G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.
FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 135 237
FT      CARBOHYD 237 251
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32525 MW; 48448F9E091E1802 CRC64;

Query Match
Best Local Similarity 87.0%; Score 1343; DB 1; Length 297;
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY      1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTLIITAAI 60
DB      1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTLIITAAI 60
QY      61 FIASANHKVLTITTAIIODATSOIKNTTPYLTQDPOLGISFSNLSSETTSQTTLASTTP 120
DB      61 FIASANHKITSTTTIIODATNOIKNTTPYLTQNPOLGISPSNPDSITSLITLIDSTTP 120
QY      121 GVKSNLOPTVTKNTNTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 GVKSTLOSTTVGKNTNTTQAOQPKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      181 TCMAICKRIPNKKPKGKTTTKPKTKPKTKKDLKQOTTKPKPEVPTTKPTPEPTINTTK 240
DB      181 TCMAICKRIPNKKPKGKTTTKPKTKPKTKKDLKQOTTKPKPEVPTTKPTPEPTINTTK 240
QY      241 TNITTTLLTNNTGNPKLTLSOMETFHSSEGNLSPSOVSTTSEHPSPSSPNTT 297
DB      241 TNITTTLLTNNTGNPKLTLSOMETFHSSEGNLSPSOVSTTSEHPSPSSPNTT 297

RESULT 4
VGLG_HRSV6
ID      VGLG_HRSV6 STANDARD: PRT; 297 AA.
AC      P27025;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-AUG-1992 (Rel. 23, Last annotation update)
DE      Major surface glycoprotein G (Attachment glycoprotein G).
GN      G.
OS      Human respiratory syncytial virus (strain rsb6256).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11256;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91374005; PubMed=1895054;
RA      Cane P.A., Matthews D.A., Pringle C.R.;
RT      "Identification of variable domains of the attachment (G) protein of
RT      subgroup A respiratory syncytial viruses.";
RL      J. Gen. Virol. 72:2091-2096(1991).
CC      -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC      RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01208; J01208.
DR      InterPro: IPR000925; Glycoprot-G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.

```

```

FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 103 103
FT      CARBOHYD 135 135
FT      CARBOHYD 237 237
FT      CARBOHYD 251 251
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

Query Match
Best Local Similarity 87.0%; Score 1343; DB 1; Length 297;
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY      1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTLIITAAI 60
DB      1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTLIITAAI 60
QY      61 FIASANHKVLTITTAIIODATSOIKNTTPYLTQDPOLGISFSNLSSETTSQTTLASTTP 120
DB      61 FIASANHKVLTITTAIIODATSOIKNTTPYLTQNPOLGISFSNLSSETTSQTTLASTTP 120
QY      121 GVKSNLOPTVTKNTNTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 SAESTPOSTVTKNTNTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      181 TCMAICKRIPNKKPKGKTTTKPKTKPKTKKDLKQOTTKPKPEVPTTKPTPEPTINTTK 240
DB      181 TCMAICKRIPNKKPKGKTTTKPKTKPKTKKDLKQOTTKPKPEVPTTKPTPEPTINTTK 240
QY      241 TNITTTLLTNNTGNPKLTLSOMETFHSSEGNLSPSOVSTTSEHPSPSSPNTT 296
DB      241 TNITTTLLTNNTGNPKLTLSOMETFHSSEGNLSPSOVSTTSEHPSPSSPNTT 296

RESULT 5
VGLG_HRSV7
ID      VGLG_HRSV7 STANDARD: PRT; 297 AA.
AC      P27026;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-AUG-1992 (Rel. 23, Last annotation update)
DE      Major surface glycoprotein G (Attachment glycoprotein G).
GN      G.
OS      Human respiratory syncytial virus (strain rsb6614).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11257;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91374005; PubMed=1895054;
RA      Cane P.A., Matthews D.A., Pringle C.R.;
RT      "Identification of variable domains of the attachment (G) protein of
RT      subgroup A respiratory syncytial viruses.";
RL      J. Gen. Virol. 72:2091-2096(1991).
CC      -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC      RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01209; J01209.
DR      InterPro: IPR000925; Glycoprot-G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.
FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 103 103
FT      CARBOHYD 135 135
FT      CARBOHYD 237 237
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

```

FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA: 32670 MW: 58384028EA37ACD CRC64;  
Query Match 86.0%; Score 1328; DB 1; Length 297;  
Best Local Similarity 87.8%; Pred. No. 2.3e-80;  
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
1 MSKTQDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
QY 121 GVSXNLOPTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
121 SAKSTPQSTTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
DB 121 SAKSTPQSTTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
QY 181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
DB 181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
QY 241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296  
241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296  
DB 241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 6  
VGLG\_HRSV4 STANDARD; PRT; 297 AA.  
ID VGLG\_HRSV4  
AC P27023;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G  
OS Human respiratory syncytial virus (strain rsb5857).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.";  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01207; J01207.  
DR InterPro: IPR000925; Glycoprot\_G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA: 32772 MW: 104886CA475936BE CRC64;

Query Match 85.8%; Score 1324; DB 1; Length 297;  
Best Local Similarity 87.8%; Pred. No. 4.2e-80;  
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
1 MSKTQDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
QY 121 GVSXNLOPTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
121 SAKSTPQSTTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
DB 121 SAKSTPQSTTVTKNTTTOPTSKPTTKORONKPNKPNNDPHEVFNFPVPCISNNP 180  
QY 181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
DB 181 TCWAICRIPNKKPKKKTTPKTKPTKTKDLKPQTPKREVPPTKPEPTINTTK 240  
QY 241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296  
241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296  
DB 241 TMTITLLTNTGNPKLTSMETFHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 7  
VGLG\_HRSV5 STANDARD; PRT; 298 AA.  
ID VGLG\_HRSV5  
AC P27024;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G  
OS Human respiratory syncytial virus (strain rsb6190).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.";  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01207; J01207.  
DR InterPro: IPR000925; Glycoprot\_G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 298 AA: 32769 MW: 4D74B54D34D7B5A5 CRC64;

Query Match 85.0%; Score 1313; DB 1; Length 298;  
Best Local Similarity 86.9%; Pred. No. 2.2e-79;  
Matches 259; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
1 MSKTQDQRTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120  
DB 61 FIASANHKVLTITAIIDATSOIKNTPTVLTQDPOLGIFSNSLSEITSOPTTALTP 120

```

Db      61 FIASANKVTLTAIIQDATSQIKNTTPYLTQNPOLGISFFNLSCSTQTAIIALTTP 120
QY      121 GYKSNLOPTTVKTKNTTTTQTOPSKPTKORONKPPNKPNDFHFEVFNVCSCISNMP 180
        121 SVESILQSTTVKTKNTTTTQIOPSKPTKORONKPPNKPNDFHFEVFNVCSCISNMP 180
Db      181 TCMAICKRIPNKKPKGKTTTKPTTKKDLKPQTTPKREVPPTKPTPEPTINTTK 240
        181 TCMAICKRIPNKKPKGKTTTKPTTKKDLKPQTTPKREVPPTKPTPEPTINTTK 240
QY      241 TTTTLLTNNTGPNKLTSSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298
        241 PNRTTLTNSTTGNLHSTSEETLHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298
Db      241 PNRTTLTNSTTGNLHSTSEETLHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 8
VGLG_HRSV2
ID VGLG_HRSV2 STANDARD: PRT: 297 AA.
AC P27021:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb642).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11252;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.;"
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR: J01204; J01204.
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32745 MW; FC22AVF3A8BEF67C CRC64;

Query Match 84.7%; Score 1307; DB 1; Length 297;
Best Local Similarity 86.5%; Pred. No. 5,4e-79;
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

```

```

QY      241 TTTTLLTNNTGPNKLTSSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTR 297
        241 TTTTLLTNNTGPNKLTSSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTR 297
Db      241 TTTTLLTNNTGPNKLTSSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTR 297

RESULT 9
VGLG_HRSV8
ID VGLG_HRSV8 STANDARD: PRT: 292 AA.
AC P23041:
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup B / strain 8/60).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90357765; PubMed=1697126;
RA Sullender W.M., Anderson K., Wertz G.W.;
RT "The respiratory syncytial virus subgroup B attachment glycoprotein:
RT analysis of sequence, expression from a recombinant vector, and
RT evaluation as an immunogen against homologous and heterologous
RT subgroup virus challenge.;"
RL Virolgy 178:195-203(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374595; PubMed=1895391;
RA Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;
RT "Genetic diversity of the attachment protein of subgroup B
RT respiratory syncytial viruses.;"
RL J. Virol. 65:5425-5434(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M55633; AAA47413.1; -
DR EMBL: M73545; AAA47408.1; -
DR PIR: A37077; MGN260.
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 292 AA; 32143 MW; BEC60C85EF057BB5 CRC64;

Query Match 49.7%; Score 767.5; DB 1; Length 292;
Best Local Similarity 55.3%; Pred. No. 1e-43;
Matches 162; Conservative 34; Mismatches 94; Indels 3; Gaps 3;

```

QY 61 FIASANKVYLITAIIDQATISQIKNTTPPYLTITOPOLGISINSLEITSOTTTLASTP 120  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |  
Db 61 FIISANKVLTLLTYTVOYQIKNHNKSTKNISTYLTQVPPERVNSSKPTTSPHTNSALISP 120  
QY 121 GVSXNLGCTTVTKNTKTMTTTOFQPSKPPTKOKHNPENKPNNDFFHEVFNFPCSGSNP 180  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |  
Db 121 NTKSETHHTTMOGRITTSQTGNKPSRKSXSKNDPKPKRDYHFVEVNFPCSIGCNG 180  
QY 181 TCMAICKRIPIKKPEKKTTPPKPPFKTT-KKDLPQTTKPREVPVTKEEPLINTT 239  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |  
Db 181 LCKSICKTIPIPNKPKPKKPIKPIKPKTKTKNNKRDPKPAKPKKEIITNTPKRPYKTT 240  
QY 240 KTNITLTLTNNTGNPKLTQSOMEFHETSESGNLSPOVSTSESHPSPSP 292  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |  
Db 241 ERDISTSOSTYIDITTPKTYITQQSLHSTSENTPSSIQITASE-PS-TSNP 291

RESULT 10  
VGLG\_HRSVL STANDARD; PRT; 292 AA.

ID VGLG\_HRSVL STANDARD; PRT; 292 AA.  
AC P20896;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-NOV-1991 (Rel. 20, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (subgroup B / strain 18537).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
CC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
RX NCBI\_TaxID=11251;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=87289657; PubMed=2441388;  
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
RT "The G glycoprotein of human respiratory syncytial viruses of  
subgroups A and B: extensive sequence divergence between  
antigenically related proteins." J. Virol. 64:5625-5629(1987).  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  
CC -1 FUNCTION: UNLIKE THE OTHER PARAMYOXOVIRUS ATTACHMENT PROTEINS, THE  
RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
HEMAGGLUTININATING ACTIVITIES.  
CC -1 SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1 PTR: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation at  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).

EMBL: M17213; AAA47412.1; .  
PIR: B32703; MGNZ18  
InterPro: IPR000925; Glycoprot\_G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 37 POTENTIAL.  
FT DOMAIN 38 66 POTENTIAL.  
FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).  
FT CARBOHD 81 81 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHD 86 86 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 292 AA; 32306 MW; BC8C59F69CAVAFC2 CRC64;

Query Match 49.3%; Score 761; DB 1; Length 292;  
Best Local Similarity 54.9%; Pred No. 2.7e-43;  
Matches 158; Conservative 35; Mismatches 93; Indels 2; Gaps 2;

1 MSKNDDQRFAKTLKRWDTLNHLFLSSGLYKLNKSVQAQILSLIAMIIISLIITTAII 60  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |  
| | | | | | | | | | | | | | | | : | | | | | | | | | | |

```

Db      1  MSKHNQTAFTLEKWTMDTLNHLIVISSCLRYLNLKSTIAOTALSVLAMIISTILLIAIT 60
Oy      61  FLASNNHVLTTLTALIDPATSOIKNTPTVLTQDPOLGISFNSLSEITQTTLLASTTP 120
        ||||||| : : : : : ||| : : : : : || : : : : :
Db      61  FLSNNHVLTTLTVVQVQIKNHTEKNISTVLTQVPPREVNSSKPTTSPHTNSATISP 120
Oy      121  GYSNNLQPTTVKTKNTTQTQTOPSKPTTKQKONPPKNPNNDHFVEFVNFVPCISGNNP 180
        || : : : : : || : : : : : || : : : : : || : : : : : ||
Db      121  NTKSTHTHTTQGTGRITTSQTQNNKPSKSKSNPKPKDDYHFEVFNFPSCISGNNQ 180
Oy      161  TCMALCKRIPLPKKPKKTKTTTPKPKPKPKT -KKDLPQRTKPREVLTPTKTEPTINT 239
        | : : : : | : : : : | : : : : | : : : : | : : : : |
Db      161  LCKSCKTTPSKPKPKPKKPKTKPKPKTKPKTKPKDKPKPKPKKEITTPAKPKPKTKT 240
Oy      240  KTNITTLTLNNTTGNPKLTQSMEFHSSTSEGNLSPOVSTSEHPS 287
        : : : : : | : : : : | : : : : | : : : : | : : : : |
Db      241  EMDTSSQSTVLDITPTPTVLTQQLSHSTSENNPSSQIPTASE-PS 287

RESULT 11
VEIGL_BRSV1 STANDARD; PRT; 263 AA.
AC      010683:
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1998 (Rel. 36, Last annotation update)
DE      Major surface glycoprotein G (Attachment glycoprotein G).
GN      G.
OS      Bovine respiratory syncytial virus (strain 127) (BRV).
OC      Viruses; ssRNA, negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_Taxid=82818;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97288324; PubMed=9143302;
RA      Furez J., Roberts S., Weitz G., Taylor G.;
RT      "Antigenically distinct G glycoproteins of BRV strains share a high
        degree of genetic homogeneity.";
RL      Virology 231:48-58(1997).
CC      -I- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
        RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
        HEMAGGLUTININATING ACTIVITIES.
CC      -I- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
        CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -I- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
        HRS VIRUS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb.ch/announce/
        or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y08716; CAAG9966.1; .
DR      HSSP; P22261; IBRV.
DR      InterPro; IPR000925; Glycoprot.G.
DR      Pfam; PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.
FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 263
FT      DISULFID 173 186
FT      DISULFID 176 182
FT      CARBOHYD 127 127
FT      CARBOHYD 163 163
FT      CARBOHYD 251 251
SQ      SEQUENCE 263 AA; 28948 MW; D592D7888989535 CRC64;
Query Match 20.0%; Score 309.5; DB 14; Length 263;
Best Local Similarity 32.5%; Pred. No. 8.3e-14;
Matches 90; Conservative 32; Mismatches 116; Indels 39; Gaps 7;

```











GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:27:43 ; Search time 35.9849 Seconds  
(without alignments)  
1706.328 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDORTAKTEKTDLT.....VSTSEHPSPSPPTNTTQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21:\*\*  
1: sp\_archaea:\*\*  
2: sp\_bacteria:\*\*  
3: sp\_fungi:\*\*  
4: sp\_human:\*\*  
5: sp\_invertebrate:\*\*  
6: sp\_mammal:\*\*  
7: sp\_mmc:\*\*  
8: sp\_organelle:\*\*  
9: sp\_phage:\*\*  
10: sp\_plant:\*\*  
11: sp\_prodent:\*\*  
12: sp\_virus:\*\*  
13: sp\_vertebrate:\*\*  
14: sp\_unclassified:\*\*  
15: sp\_virus:\*\*  
16: sp\_bacteriap:\*\*  
17: sp\_archaeap:\*\*

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537	99.5	298	12	001929 human respi
2	1533	99.3	298	12	082068 human respi
3	1393	90.2	298	12	009634 human respi
4	1389	90.0	298	12	009719 respiratory
5	1365	88.4	297	12	09YVB5 human respi
6	1364	88.3	297	12	082066 human respi
7	1356	87.8	298	12	09YVB4 human respi
8	1354	87.7	297	12	082071 human respi
9	1345	87.1	297	12	082067 human respi
10	1344	87.0	297	12	082058 human respi
11	1339	86.7	297	12	091947 human respi
12	1336	86.5	293	12	09YVB3 human respi
13	1334	86.4	295	12	086359 respiratory
14	1334	86.4	295	12	082065 respiratory
15	1334	86.4	298	12	082065 human respi
16	1331	86.2	297	12	082074 human respi

17	1331	86.2	297	12	082064 human respi
18	1328	86.0	297	12	082063 human respi
19	1326	85.9	292	12	09YVB2 human respi
20	1326	85.9	292	12	082077 human respi
21	1325	85.8	298	12	082060 human respi
22	1324	85.8	297	12	091946 human respi
23	1320	85.5	297	12	082070 human respi
24	1319	85.4	293	12	09YVB1 human respi
25	1319	85.4	295	12	086361 human respi
26	1319	85.4	297	12	082079 human respi
27	1319	85.4	297	12	082057 human respi
28	1318	85.4	297	12	09YVC8 human respi
29	1317	85.3	298	12	082075 human respi
30	1317	85.3	298	12	082062 human respi
31	1317	85.3	298	12	082056 human respi
32	1316	85.2	297	12	082072 human respi
33	1315	85.2	295	12	086357 respiratory
34	1314	85.1	297	12	082078 human respi
35	1314	85.1	297	12	091944 human respi
36	1314	85.1	298	12	082073 human respi
37	1314	85.1	298	12	082076 human respi
38	1312	85.0	298	12	011978 human respi
39	1311	84.9	292	12	09YVB0 human respi
40	1311	84.9	298	12	091945 human respi
41	1308	84.7	296	12	086358 respiratory
42	1306	84.6	298	12	082069 human respi
43	1305	84.5	298	12	082061 human respi
44	1298	84.1	295	12	086356 respiratory
45	1282	83.0	279	12	090678 human respi

#### ALIGNMENTS

RESULT 1  
ID 001929 PRELIMINARY; PRT; 298 AA.  
AC 001929;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (subgroup A / strain Long).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9106351; PubMed=2249671;  
RA Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;  
RT "Frame shift mutations as a novel mechanism for the generation of  
RT neutralization resistant mutants of human respiratory syncytial  
RT virus.";  
RL EMBL J. 9:4181-4187(1990).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: THIS PROTEIN MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE  
CC CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
CC EMBL: X17085; CAA34937.1; -;  
DR InterPro: IPR000925; Glycoprot G.  
DR InterPro: IPR003880; Prntine.attach.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 INTRACELLULAR (BY SIMILARITY).  
FT TRANSMEM 38 66  
FT DOMAIN 67 298 EXTRACELLULAR (BY SIMILARITY).  
FT DOMAIN 205 211 ESSENTIAL FOR EPITOPE 636 INTEGRITY.  
FT SITE 64 65 CLEAVAGE (BY PROTEASE) (POTENTIAL).

FT SITE 74 75 CLEAVAGE (BY PROTEASE) (POTENTIAL)  
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL)  
 SQ SEQUENCE 298 AA; 32771 MW; F8CD4213D97C2952 CRC64;

Query Match 99.5%; Score 1537; DB 12; Length 298;  
 Best Local Similarity 99.3%; Pred. No. 4, 5e-101;  
 Matches 296; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAIT 60  
 DB 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAIT 60  
 QY 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 DB 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 QY 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 DB 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 QY 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240  
 DB 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240  
 QY 241 TTTTTLTNNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298  
 DB 241 TTTTTLTNNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298

## RESULT 2

ID 082068 PRELIMINARY; PRT; 298 AA.  
 AC 082068;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
 DE Glycoprotein (Fragment).  
 GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBL\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SUBGROUP A;  
 RX MEDLINE=9435057; PubMed=8057427;  
 RA Garcia O., Martin M., Dopazo J., Arbliza J., Fabrales S., Russi J.,  
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
 RA Melero J.A.;  
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein."  
 RT J. VIROL. 68:3448-3459(1994).  
 RL EMBL: Z33429; CAA83872.1; -  
 DR InterPro: IPR000925; Glycoprot.G.  
 DR InterPro: IPR003880; Ppantne.attach.  
 DR Pfam: PF00802; Glycoprotein\_G.1.  
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 99.3%; Score 1533; DB 12; Length 298;  
 Best Local Similarity 99.0%; Pred. No. 8, 6e-101;  
 Matches 295; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAIT 60  
 DB 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAIT 60

QY 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 DB 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 QY 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 DB 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 QY 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240  
 DB 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240  
 QY 241 TTTTTLTNNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298  
 DB 241 TTTTTLTNNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298

## RESULT 3

ID 009634 PRELIMINARY; PRT; 298 AA.  
 AC 009634;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-JUL-1997 (TREMblrel. 04, last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
 DE Attachment protein (G).  
 GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBL\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S2;  
 RX MEDLINE=97185152; PubMed=9032893;  
 RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,  
 RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,  
 RA Pringle C.R.;  
 RT "Identification of mutations contributing to the reduced virulence of a modified strain of respiratory syncytial virus."  
 RT Vaccine 14:1637-1646(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S2;  
 RA Easton A.J.;  
 RX Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U39662; AAC57026.1; -  
 DR InterPro: IPR000925; Glycoprot.G.  
 DR InterPro: IPR003880; Ppantne.attach.  
 DR Pfam: PF00802; Glycoprotein\_G.1.  
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN\_1.  
 SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match 90.2%; Score 1393; DB 12; Length 298;  
 Best Local Similarity 91.3%; Pred. No. 6, 5e-91;  
 Matches 272; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAIT 60  
 DB 1 MSKNKQDRTAKLEKTWDTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAIT 60  
 QY 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 DB 61 FIASANHKVLTITAIODATSOIKNTPTTYLTODPOLGISFSNLSITSOITTLASTP 120  
 QY 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 DB 121 GVSXNLOPTVTKNTTTTQTOPSKPTTKORONKPPKNPNNDHFVFNVPSCISNNP 180  
 QY 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240  
 DB 181 TCMAICKRIPNKKRPGKTTTKPTKPKTKDLKQPTTKPKREVPPTKPEEPTINTTK 240



DR InterPro: IPR003880; Ppantne\_attach.  
 DR Pfam: PF00802; Glycoprotein\_G.1.  
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
 SO SEQUENCE 297 AA; 32589 MW; F210F6C68193F5DD CRC64;

Query Match 88.3%; Score 1364; DB 12; Length 297;  
 Best Local Similarity 89.9%; Pred. No. 7.2e-89;  
 Matches 267; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 Db 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 QY 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 Db 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 QY 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180  
 Db 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180  
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 QY 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 297  
 Db 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 297

## RESULT 7

QY YB4 PRELIMINARY; PRT; 298 AA.

AC 09YB4;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 OS Attachment glycoprotein G.  
 DE Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WV5222;  
 RX MEDLINE=99022964; PubMed=9806017;  
 RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.,  
 RT "Antigenic and genetic diversity among the attachment proteins of  
 RT group A respiratory syncytial viruses that have caused repeat  
 RT infections in children."  
 RL J. Infect. Dis. 178:925-932(1998).  
 DR EMBL: AF065406; AAD02942.1; -  
 DR InterPro: IPR000925; Glycoprot\_G.  
 DR InterPro: IPR003880; Ppantne\_attach.  
 DR Pfam: PF00802; Glycoprotein\_G.1.  
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
 SO SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 87.8%; Score 1356; DB 12; Length 298;  
 Best Local Similarity 88.9%; Pred. No. 2.6e-88;  
 Matches 265; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 Db 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 QY 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 Db 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 QY 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180  
 Db 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180

QY 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 QY 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 298  
 Db 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 298

## RESULT 8

QY 082071 PRELIMINARY; PRT; 297 AA.

AC 082071;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Glycoprotein.  
 GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SUBGROUP A;  
 RX MEDLINE=94335057; PubMed=8057427;  
 RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisse S., Russi J.,  
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
 RA Melero J.A.;  
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
 RT A): cocirculating lineages and correlation of genetic and antigenic  
 RT changes in the G glycoprotein."  
 RL J. Virol. 68:5448-5459(1994).  
 DR EMBL: Z33432; CAA83875.1; -  
 DR InterPro: IPR000925; Glycoprot\_G.  
 DR InterPro: IPR003880; Ppantne\_attach.  
 DR Pfam: PF00802; Glycoprotein\_G.1.  
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
 SO SEQUENCE 297 AA; 32581 MW; AD947B2318193306 CRC64;

Query Match 87.7%; Score 1354; DB 12; Length 297;  
 Best Local Similarity 89.2%; Pred. No. 3.6e-88;  
 Matches 265; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 Db 1 MSKNKDDORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60  
 QY 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 Db 61 FIASANHKVLTITAIQDATSQIKNTTPYVLQDPOLGIFSNSLSEITTSQTTIIASTP 120  
 QY 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180  
 Db 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPKNPNDFEFVFNVPSCISNNP 180  
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPKTKKDKLPOTTKPEVPTTKPTEPTTINTK 240  
 QY 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 297  
 Db 241 TITTTLLTNNLTGNPKLTSSQMETFHSSTSEGNLSPSOVSTTSEHPSPSPPTTR 297

## RESULT 9

QY 082067 PRELIMINARY; PRT; 297 AA.

AC 082067;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Glycoprotein.  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SUBGROUP A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabraile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): co-circulating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:548-549(1994).  
DR EMBL: Z33428; CA83871.1;  
DR InterPro: IPR000925; Glycoprot\_G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.  
SQ SEQUENCE 297 AA; 32555 MW; 9D0A69BECAB6C8 CRC64;

Query Match 87.1%; Score 1345; DB 12; Length 297;  
Best Local Similarity 88.9%; Pred. No. 1.6e-87;  
Matches 264; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
DB 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
OY 61 FIASANKVLTTLTALIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120  
DB 61 FIASANKVLTTLTALIODATNOIKNTPTTYLQDPOLGISFNSLSEITSLITLIDSTTP 120  
OY 121 GVSNSLOPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180  
DB 121 GVSNSLOPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180  
OY 181 TCMAICKRIPKKRGKKTTPPTKPKPTTKKDKLPOTKPKREVPPTKPEEPTINTTK 240  
DB 181 TCMAICKRIPKKRGKKTTPPTKPKPTTKKDKLPOTKPKREVPPTKPEEPTINTTK 240  
OY 241 TNIITTLTNTTGNPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPPTNTT 297  
DB 241 TNIITTLTNTTGNPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPPTNTT 297

RESULT 10  
082058 PRELIMINARY; PRT; 297 AA.  
AC 082058;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE (MAD-1-89) subgroup A, G glycoprotein.  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SUBGROUP A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabraile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): co-circulating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:548-549(1994).

DR EMBL: Z33456; CA83879.1;  
DR InterPro: IPR000925; Glycoprot\_G.  
DR InterPro: IPR003880; Pantine\_attach.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.  
SQ SEQUENCE 297 AA; 32741 MW; 0E567A174BE64964 CRC64;

Query Match 87.0%; Score 1344; DB 12; Length 297;  
Best Local Similarity 88.5%; Pred. No. 1.8e-87;  
Matches 262; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
DB 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
OY 61 FIASANKVLTTLTALIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120  
DB 61 FIASANKVLTTLTALIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120  
OY 121 GVSNSLOPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180  
DB 121 GVSNSLOPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180  
OY 241 TNIITTLTNTTGNPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPPTNTT 296  
DB 241 TNIITTLTNTTGNPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPPTNTT 296

RESULT 11  
091947 PRELIMINARY; PRT; 297 AA.  
AC 091947;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Attachment protein.  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CH34;  
RX MEDLINE=98418504; PubMed=9747732;  
RA Peret T.C.T., Hall C.B., Schenabel K.C., Golub J.A., Anderson L.J.;  
RT "Circulation Patterns of Genetically Distinct Group A and B Strains of  
RT Human Respiratory Syncytial Virus in a Community.";  
RL J. Gen. Virol. 79:2221-2229(1998).  
DR EMBL: AF065257; AAC36327.1;  
DR InterPro: IPR000925; Glycoprot\_G.  
DR InterPro: IPR003880; Pantine\_attach.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.  
SQ SEQUENCE 297 AA; 32564 MW; 377C7C8CDD0A3645 CRC64;

Query Match 86.7%; Score 1339; DB 12; Length 297;  
Best Local Similarity 88.2%; Pred. No. 4.2e-87;  
Matches 262; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
DB 1 MSKNDQRTAKTLEKWTWTLNHLFTSSGLYKLNKSAVQITSLIAMISTSLITTAI 60  
OY 61 FIASANKVLTTLTALIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120  
DB 61 FIASANKVLTTLTALIODATNOIKNTPTTYLQDPOLGISFNSLSEITSLITLIDSTTP 120  
OY 121 GVSNSLOPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180

```
DB 121 GVKSTLSTSTVGKTKNTTQTAQPKPTTKORQNTPSKPNNDHFEVFNVPSCISNNP 180
OY 181 TCMAICRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTK 240
DB 181 TCMAICRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTK 240
OY 241 TMTTLLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTR 297
DB 241 TMTTLLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTR 297

RESULT 12
OYVVB3 PRELIMINARY: PRT: 293 AA.
ID OYVVB3
AC OYVVB3;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Attachment glycoprotein G (Fragment).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB973;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065407; AAD02943.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00802; Glycoprotein_G. 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 293 AA; 32125 MW; 17B5B43396A63CFC CRC64;
SQ

Query Match 86.5%; Score 1336; DB 12; Length 293;
Best Local Similarity 90.1%; Pred. No. 6.7e-87;
Matches 264; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

OY 6 DQRTAKTEKTWDTLNHLFISSGLYKLNKSAQAQITLSIAMIISTSLITAIIFASA 65
DB 1 DQRTAKTEKTWDTLNHLFISSGLYKLNKSAQAQITLSIAMIISTSLITAIIFASA 60
OY 66 NHKVTLLTAIIODATSQIKNTPTLYLTOPOLGISFSNLSSETTSQTTILASTPVGKSN 125
DB 61 NHKVTLLTAIIODATSQIKNTPTLYLTOPOLGISFSNLSSETTSQTTILASTPVGKSN 120
OY 126 LQPTTVKTKNTTQTPQSKPTTKORQNPKNPNNDHFEVFNVPSCISNNPTCAI 185
DB 121 LQPTTVKTKNTTQTPQSKPTTKORQNPKNPNNDHFEVFNVPSCISNNPTCAI 180
OY 186 CKRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTKNTT 245
DB 181 CKRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTKNTT 240
OY 246 TLTNTTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTRQ 298
DB 241 TLTNTTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTRQ 293

RESULT 13
OYVVB3 PRELIMINARY: PRT: 295 AA.
ID OYVVB3;
AC OYVVB3;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
```

```
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE G protein (Fragment).
OS respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX Cane P.A.;
RA Submitted (Apr-1993) to the EMBL/Genbank/DBJ databases.
RL EMBL: X73353; CAA51764.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G. 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;
SQ

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.4%; Pred. No. 9.3e-87;
Matches 260; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

OY 3 KNDQRTAKTEKTWDTLNHLFISSGLYKLNKSAQAQITLSIAMIISTSLITAIIFI 62
DB 1 KNDQRTAKTEKTWDTLNHLFISSGLYKLNKSAQAQITLSIAMIISTSLITAIIFI 60
OY 63 ASANHKVTLTAIIODATSQIKNTPTLYLTOPOLGISFSNLSSETTSQTTILASTPVG 122
DB 61 ASANHKVTLTAIIODATSQIKNTPTLYLTOPOLGISFSNLSSETTSQTTILASTPVG 120
OY 123 KSNLQPTTVKTKNTTQTPQSKPTTKORQNPKNPNNDHFEVFNVPSCISNNPTC 182
DB 121 ESTQSTTVKTKNTTQTPQSKPTTKORQNPKNPNNDHFEVFNVPSCISNNPTC 180
OY 183 WAIICKRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTKTN 242
DB 181 WAIICKRIPKPKGKTKTKPTKPKTKKDLKDPQTKRPEVPTKPEPTINTTKTN 240
OY 243 ITTLLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTR 296
DB 241 ITTLLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPQSSPPTTR 294

RESULT 14
OYVVB3 PRELIMINARY: PRT: 295 AA.
ID OYVVB3;
AC OYVVB3;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE G protein (Fragment).
OS respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
```

[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN-RSB89-642;  
RC Cane P.A.;  
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL; XJ3334; CAAS1765.1; -.  
DR InterPro; IPR000925; Glycoprot\_G.  
DR Pfam; PF00802; Glycoprotein.G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
FT NON\_TER 1  
SQ SEQUENCE 295 AA; 32306 MW; 671DCF041570284 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;  
Best Local Similarity 88.2%; Pred. No. 9,3e-87;  
Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNDQRTAKLEKTWDTLNLHLFISGLYKLNKSAQITLSILAMISTSLITAIIFI 62  
DB 1 KNDQRTAKLEKTWDTLNLHLFISGLYKLNKSAQITLSILAMISTSLITAIIFI 60  
QY 63 ASANHKVTLTTATIODATSOIKNTPTTYLTODPOLGISNSLSEITTSQTTTILASTTPGV 122  
DB 61 ASANHKITSTTTIIQDAIWOIKNTPTTYLTQNPOLGISNSBDSITSLITLIDSTTPGV 120  
QY 123 KSNLQPTTVKTKNTTTQTQPSKPTTKORONKPPNNDPHEVEFNFPVCSISNNPTC 182  
DB 121 KSTLGSTVGTKNWTTTQOQPNKPTTKORONKPPNNDPHEVEFNFPVCSISNNPTC 180  
QY 183 WAICCKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPKVPTTKPEEPTINTTKTN 242  
DB 181 WAICCKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPKVPTTKPEEPTINTTKTN 240  
QY 243 ITTTLTNTGTPKLTSGMETFHSSEGNLSPSCVSTTSSEHPSQSPSPNTTR 297  
DB 241 ITTTLTNTGTPKLTSGMETFHSSEGNLSPSCVSTTSSEHPSQSPSPNTTR 295

RESULT 15

ID 082065 PRELIMINARY; PRT: 298 AA.  
AC 082065;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Glycoprotein.  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBGROUP A;  
RX MEDLINE=9435057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Weier J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL; Z33426; CAA83869.1; -.  
DR InterPro; IPR000925; Glycoprot\_G.  
DR Pfam; PF00802; Glycoprotein.G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
SQ SEQUENCE 298 AA; 32729 MW; 4E890FPD43B45744 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 298;  
Best Local Similarity 88.2%; Pred. No. 9,4e-87;  
Matches 261; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSKNQDRTAKLEKTWDTLNLHLFISGLYKLNKSAQITLSILAMISTSLITAIIFI 60  
DB 1 MSKNQDRTAKLEKTWDTLNLHLFISGLYKLNKSAQITLSILAMISTSLITAIIFI 60  
QY 61 FIASANHKVTLTTATIODATSOIKNTPTTYLTODPOLGISNSLSEITTSQTTTILASTTP 120  
DB 61 FIASANHKVTLTTATIODATSOIKNTPTTYLTQNPOLGISNSBDSITSLITLIDSTTPGV 120  
QY 121 GVSMLQPTTVKTKNTTTQTQPSKPTTKORONKPPNNDPHEVEFNFPVCSISNNPTC 180  
DB 121 SAESTPLSTTVKTKNTTTQOQPNKPTTKORONKPPNNDPHEVEFNFPVCSISNNPTC 180  
QY 181 TCWAICCKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPKVPTTKPEEPTINTTKTN 240  
DB 181 TCWAICCKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPKVPTTKPEEPTINTTKTN 240  
QY 241 ITTTLTNTGTPKLTSGMETFHSSEGNLSPSCVSTTSSEHPSQSPSPNTTR 296  
DB 241 ITTTLTNTGTPKLTSGMETFHSSEGNLSPSCVSTTSSEHPSQSPSPNTTR 296

Search completed: May 1, 2003, 13:46:23  
Job time : 36.9849 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:55:01 ; Search time 1442.68 Seconds

(without alignments)  
14423.488 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715  
Sequence: 1 caacaagtcacatacaaac.....gtcgtatttaaaaaaaaaa 715

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	920	6 A16257	A16257 Synthetic R
2	715	100.0	920	6 A16258	A16258 Synthetic R
3	715	100.0	920	6 AR080406	AR080406 Sequence
4	715	100.0	920	6 AR092530	AR092530 Sequence
5	715	100.0	920	6 AR122885	AR122885 Sequence
6	715	100.0	920	6 AR123540	AR123540 Sequence
7	715	100.0	920	6 AR148357	AR148357 Sequence
8	710	99.3	923	6 HRSRMAG	X17085 Human resp
9	709	99.2	922	14 HRSVGL16	Z33429 Human resp
10	699.2	97.8	917	14 RSHGLYG	M17212 Human resp
11	696	97.3	894	6 AR080424	AR080424 Sequence
12	696	97.3	894	6 AR092548	AR092548 Sequence
13	654.8	91.6	8510	14 RSHICE	M11466 Human resp
14	654.8	91.6	15222	6 AR093219	AR093219 Sequence
15	654.8	91.6	15222	14 HRU50362	U50362 Human resp
16	654.8	91.6	15222	14 HRU50363	U50363 Human resp
17	654.8	91.6	15222	14 HRU63644	U63644 Human resp
18	654.8	91.6	15222	14 RSHSPD	M74568 Human resp
19	654.8	91.6	15223	6 AR089137	AR089137 Sequence
20	654.8	91.6	15223	14 AF035006	AF035006 Human res
21	649.6	90.9	914	14 PARSENG	X03149 Respiratory
22	649.6	90.9	918	6 AX047075	AX047075 Sequence
23	646.2	90.4	897	6 AX339011	AX339011 Sequence
24	641.8	89.8	15190	14 HRU39662	HRU39662 Human resp
25	641.8	89.8	15191	14 RSU39661	RSU39661 Respiratory
26	628.6	87.9	840	14 AF193308	AF193308 Human res
27	628.2	87.9	918	14 AF065406	AF065406 Human res
28	625.4	87.5	840	14 AF193306	AF193306 Human res
29	621	86.9	922	14 HRSVGL3	Z33412 Human resp
30	621	86.9	922	14 HRSVGL1	Z33416 Human resp
31	621	86.9	922	14 HRSVGL11	Z33417 Human resp
32	621	86.9	922	14 HRSVGL12	Z33418 Human resp
33	620.8	86.8	887	14 AF065409	AF065409 Human res
34	620.2	86.7	918	14 AF065405	AF065405 Human res
35	619.4	86.6	922	14 HRSVGL2	Z33415 Human resp
36	618.4	86.5	922	14 HRSVGL1	Z33410 Human resp
37	618.4	86.5	922	14 HRSVGL1	Z33410 Human resp
38	618.4	86.5	922	14 HRSVGL12	Z33410 Human resp
39	617.8	86.4	922	14 AF065256	AF065256 Human res
40	617.8	86.4	922	14 AF065258	AF065258 Human res
41	617.8	86.4	922	14 HRSVGL13	Z33426 Human resp
42	617.8	86.4	922	14 HRSVGL9	Z33422 Human resp
43	616.8	86.3	922	14 HRSVGLX1	Z33493 Human resp
44	616.2	86.2	922	14 HRSVGL3	Z33416 Human resp
45	616.2	86.2	922	14 HRSVGL4	Z33417 Human resp

#### ALIGNMENTS

RESULT 1  
LOCUS A16257 920 bp DNA  
DEFINITION Synthetic RSV G gene (seq ID No: 7).  
ACCESSION A16257  
VERSION A16257.1 GI:640933  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
artificial sequences.  
FEATURES  
source 1..920  
/organism="synthetic construct"  
/db\_xref="taxon:32650"  
8..904  
gene



SOURCE Unknown.  
ORGANISM Unknown.  
KEYWORDS Unclassified.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 5968776-A 7 19-OCT-1999;  
FEATURES location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.6e-149;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAGTCACTAATCACTGCAATCATACAGATGCAACAGCCGATCAAGACACA 60  
Db CACAAGTCACTAATCACTGCAATCATACAGATGCAACAGCCGATCAAGACACA 265  
QY 61 ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 120  
Db ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 325  
QY 121 ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 180  
Db ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
QY 326 ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
Db ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 240  
Db CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
QY 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
Db CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
QY 241 ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 300  
Db ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 505  
QY 446 ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 300  
Db ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 505  
QY 301 GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 360  
Db GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
QY 506 GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
Db GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Db AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Db ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
QY 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
Db ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
QY 481 ACCACCAAGCCCAAGAGAGCCCAACCATCAACACACCAACCAACCAACCAACCAACCAACCA 540  
Db ACCACCAAGCCCAAGAGAGCCCAACCATCAACACACCAACCAACCAACCAACCAACCAACCA 745  
QY 541 CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 600  
Db CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 805  
QY 601 TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 660  
Db TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
QY 806 TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
Db TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
QY 661 TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 715  
Db TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
QY 866 TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
Db TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
RESULT 4  
AR092530 920 bp DNA linear PAT 08-SEP-2000  
LOCUS AR092530  
DEFINITION Sequence 7 from patent US 599169.  
ACCESSION AR092530

VERSION AR092530.1 GI:10019284  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 5998169-A 7 07-DEC-1999;  
FEATURES location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.6e-149;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAGTCACTAATCACTGCAATCATACAGATGCAACAGCCGATCAAGACACA 60  
Db CACAAGTCACTAATCACTGCAATCATACAGATGCAACAGCCGATCAAGACACA 265  
QY 61 ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 120  
Db ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 325  
QY 266 ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 325  
Db ACCCAACATACCTCTACTAGAGTCTCTAGCTTGGAATCAGCTTCTCCAACTGCTGAA 325  
QY 121 ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 180  
Db ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
QY 326 ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
Db ATTACATCAACAAGCCACCATCTACTAGTTCAACACACAGAGATCAAGTCAACCTG 385  
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 240  
Db CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
QY 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
Db CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445  
QY 241 ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 300  
Db ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 505  
QY 446 ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 300  
Db ACTCAAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 505  
QY 301 GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 360  
Db GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
QY 506 GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
Db GTGTTTAACCTTTGACCTGAGCATATGAGCAACCAACCAACCTGCTGGGCTATCTGC 565  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Db AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Db ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
QY 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
Db ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
QY 481 ACCACCAAGCCCAAGAGAGCCCAACCATCAACACACCAACCAACCAACCAACCAACCAACCA 540  
Db ACCACCAAGCCCAAGAGAGCCCAACCATCAACACACCAACCAACCAACCAACCAACCAACCA 745  
QY 541 CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 600  
Db CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 805  
QY 746 CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 805  
Db CTGCTCAACAACAACACAGAGAAATCCAAAATCAACAAGTCAAAATGGAACCTTCCAC 805  
QY 601 TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 660  
Db TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
QY 806 TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
Db TCACACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAAACATCCGACACCA 865  
QY 661 TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 715  
Db TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
QY 866 TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
Db TCACACCTCTCTACACCAACACACACAGGCAAGTATTTAAAAA 920  
RESULT 5  
AR122885 920 bp DNA linear PAT 16-MAY-2001  
LOCUS AR122885

DEFINITION Sequence 7 from patent US 6168786.  
ACCESSION ARI22885  
VERSION ARI22885.1 GI:14107851  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszshyn, M.E.  
TITLE Chimeric immunogens  
JOURNAL Patent: US 6168786-A 7 02-JAN-2001;  
FEATURES  
Location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.6e-149;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAATCACTGCAATCATACAGATGCAAGCAAGCCAGATCAAGAACACA 60  
DB 206 CACAAGTCACACTAATCACTGCAATCATACAGATGCAAGCAAGCCAGATCAAGAACACA 265  
QY 61 ACCCAACATACCTCACTCAGATCTCAGCTTGGAATCAGCTTCCAAATCTGTCTGAA 120  
DB 266 ACCCAACATACCTCACTCAGATCTCAGCTTGGAATCAGCTTCCAAATCTGTCTGAA 325  
QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 180  
DB 326 ATTACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 385  
QY 181 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAAGCC 240  
DB 386 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAAGCC 445  
QY 241 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 446 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
QY 301 GTGTTTAACCTTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 360  
DB 506 GTGTTTAACCTTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 565  
QY 361 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAAGCCCTCAAAAAACCA 420  
DB 566 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAAGCCCTCAAAAAACCA 625  
QY 421 ACCTTCAAGCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAACTACCC 480  
DB 626 ACCTTCAAGCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAACTACCC 685  
QY 481 ACCACCAAGCCCAAGAGGCCAACCATTCACACCAACCAAAACCAACCAACCAACCA 540  
DB 686 ACCACCAAGCCCAAGAGGCCAACCATTCACACCAACCAAAACCAACCAACCAACCA 745  
QY 541 CTGCTCACAACAACACACAGGAATCCAAAACCTCACAAGTCAATGGAACCTTCCAC 600  
DB 746 CTGCTCACAACAACACACAGGAATCCAAAACCTCACAAGTCAATGGAACCTTCCAC 805  
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCCACAACATCCGAGCACCA 660  
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCCACAACATCCGAGCACCA 865  
QY 661 TCACAACCTCATCTCCACCCACAACACAGCCAGTAGTTATTAATAAAAAA 715  
DB 866 TCACAACCTCATCTCCACCCACAACACAGCCAGTAGTTATTAATAAAAAA 920

RESULT 6  
ARI23540  
LOCUS ARI23540 920 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6171783.  
ACCESSION ARI23540  
VERSION ARI23540.1 GI:14108901  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszshyn, M.E.  
TITLE Infection detection method using chimeric protein  
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;  
FEATURES  
Location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.6e-149;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAATCACTGCAATCATACAGATGCAAGCAAGCCAGATCAAGAACACA 60  
DB 206 CACAAGTCACACTAATCACTGCAATCATACAGATGCAAGCAAGCCAGATCAAGAACACA 265  
QY 61 ACCCAACATACCTCACTCAGATCTCAGCTTGGAATCAGCTTCCAAATCTGTCTGAA 120  
DB 266 ACCCAACATACCTCACTCAGATCTCAGCTTGGAATCAGCTTCCAAATCTGTCTGAA 325  
QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 180  
DB 326 ATTACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 385  
QY 181 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAAGCC 240  
DB 386 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAAGCC 445  
QY 241 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 446 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
QY 301 GTGTTTAACCTTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 360  
DB 506 GTGTTTAACCTTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 565  
QY 361 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAAGCCCTCAAAAAACCA 420  
DB 566 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAAGCCCTCAAAAAACCA 625  
QY 421 ACCTTCAAGCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAACTACCC 480  
DB 626 ACCTTCAAGCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAACTACCC 685  
QY 481 ACCACCAAGCCCAAGAGGCCAACCATTCACACCAACCAAAACCAACCAACCAACCA 540  
DB 686 ACCACCAAGCCCAAGAGGCCAACCATTCACACCAACCAAAACCAACCAACCAACCA 745  
QY 541 CTGCTCACAACAACACACAGGAATCCAAAACCTCACAAGTCAATGGAACCTTCCAC 600  
DB 746 CTGCTCACAACAACACACAGGAATCCAAAACCTCACAAGTCAATGGAACCTTCCAC 805  
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCCACAACATCCGAGCACCA 660  
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCCACAACATCCGAGCACCA 865  
QY 661 TCACAACCTCATCTCCACCCACAACACAGCCAGTAGTTATTAATAAAAAA 715  
DB 866 TCACAACCTCATCTCCACCCACAACACAGCCAGTAGTTATTAATAAAAAA 920

RESULT 7  
ARI48357  
LOCUS ARI48357 920 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 7 from patent US 6225091.  
LOCUS ARI48357  
ACCESSION ARI48357.1 GI:15112447  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Sequence 7 from patent US 6225091.  
ARI48357.1  
GI:15112447  
Unknown.  
Unclassified.  
1 (bases 1 to 920)  
Klein, M.H., Du, R.-P. and Ewasysbyn, M.E.  
Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
Patent: US 6225091-A 7 01-MAY-2001;  
Location/Qualifiers  
1..920  
/organism="unknown"

380 a 290 c 95 g 155 t

Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.6e-149;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACAAGTGCATATGCAAGATGCAAGCCAGATCAAGAACA 60  
DB 206 CACAAAGTACACTAACAAGTGCATATGCAAGATGCAAGCCAGATCAAGAACA 265  
QY 61 ACCCCAACTACCTCAGTCTAGGATCTGAGCTTGAATGATCTGCAATGCTGAA 120  
DB 266 ACCCCAACTACCTCAGTCTAGGATCTGAGCTTGAATGATCTGCAATGCTGAA 325  
QY 121 ATTACATCAAAACCAACCACTACTAGCTTCACAGACCAAGAGTCAAGTCAAACTG 180  
DB 326 ATTACATCAAAACCAACCACTACTAGCTTCACAGACCAAGAGTCAAGTCAAACTG 385  
QY 181 CAACCCACACAGTCAAGCTAAAAACACACACACACCAACCCCAACCCAGCAAGCC 240  
DB 386 CAACCCACACAGTCAAGCTAAAAACACACACACACCAACCCCAACCCAGCAAGCC 445  
QY 241 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 446 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 360  
DB 506 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 565  
QY 361 AAAAAGATACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420  
DB 566 AAAAAGATACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 625  
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 626 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
QY 481 ACCACCAAGCCCAAGAGCAACATCAACACCAACCAACCAACCAACCAACCAACCA 540  
DB 686 ACCACCAAGCCCAAGAGCAACATCAACACCAACCAACCAACCAACCAACCAACCA 745  
QY 541 CTGCTCACCACACACACAGGAAATGCAAAATGCAAGTCAAAATGGAACCTTCAC 600  
DB 746 CTGCTCACCACACACAGGAAATGCAAAATGCAAGTCAAAATGGAACCTTCAC 805  
QY 601 TCAACCTCTCGAAGGCAATGTAAGCTTCTCAAGTCTCCACAACTCGAGACCA 660  
DB 806 TCAACCTCTCGAAGGCAATGTAAGCTTCTCAAGTCTCCACAACTCGAGACCA 865  
QY 661 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 715  
DB 866 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

HRSRNAG  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
FEATURES  
CDS

HRSRNAG  
Human respiratory syncytial (RS) virus mRNA for G protein.  
X17085  
GI:60306  
g protein.  
Human respiratory syncytial virus.  
Human respiratory syncytial virus.  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
1 (bases 1 to 923)  
Garcia-Barreno, B., Portela, A., Delgado, T., Lopez, J.A. and Melero, J.A.  
Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus  
EMBO J. 9 (12), 4181-4187 (1990)  
91065351  
2249671  
Location/Qualifiers  
1..923  
/organism="Human respiratory syncytial virus;"  
/db\_xref="taxon:11250"  
16..912  
/codon\_start=1  
/product="G protein"  
/protein\_id="CAA34937.1"  
/db\_xref="GI:60307"  
/translation="MSKNDQRTAKLTLEKTDWTLNLYLFISGKYLNKLSIAQITLS  
ILMIISTSLIIFALIFIASANRVTLLTAI10DASQIKNTFTVLPDOLGISFS  
NLSEITSOITLILASTPGVSKSLQPTVTKMTTTOPTSPPTKORONPKPN  
NDHPFVFNVPSCISNNPTCAICRIKPNKPKKTKTKPKRFTKTKDKLP  
TKPKRFTKTKDKLP  
SVSTTSEHSDSPSPNTTTO"

BASE COUNT 379 a 290 c 98 g 156 t

Query Match 99.3%; Score 710; DB 14; Length 923;  
Best Local Similarity 100.0%; Pred. No. 2e-148;  
Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACAAGTGCATATGCAAGATGCAAGCCAGATCAAGAACA 60  
DB 214 CACAAAGTACACTAACAAGTGCATATGCAAGATGCAAGCCAGATCAAGAACA 273  
QY 61 ACCCCAACTACCTCAGTCTAGGATCTGAGCTTGAATGATCTGCAATGCTGAA 120  
DB 274 ACCCCAACTACCTCAGTCTAGGATCTGAGCTTGAATGATCTGCAATGCTGAA 333  
QY 121 ATTACATCAAAACCAACCACTACTAGCTTCACAGACCAAGAGTCAAGTCAAACTG 180  
DB 334 ATTACATCAAAACCAACCACTACTAGCTTCACAGACCAAGAGTCAAGTCAAACTG 393  
QY 181 CAACCCACACAGTCAAGCTAAAAACCAACCAACCAACCAACCAACCAACCAACCA 240  
DB 394 CAACCCACACAGTCAAGCTAAAAACCAACCAACCAACCAACCAACCAACCAACCA 453  
QY 241 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 454 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513  
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 360  
DB 514 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 573  
QY 361 AAAAAGATACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420  
DB 574 AAAAAGATACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 633  
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 634 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693

QY 481 ACCACCAAGCCACAGAGAGCAACCATCAACACCAACCAACCAACATCAACTACA 540  
DB 694 ACCACCAAGCCACAGAGAGCCACCATCAACACCAACCAACATCAACTACA 753  
QY 541 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 600  
DB 754 CTGCTCAGCAACACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 813  
QY 601 TCACACTCTCCGAGAGCAATCAAGCCTTCTCAAGTCTCCAAACATCCAGACCCA 660  
DB 814 TCACACTCTCCGAGAGCAATCAAGCCTTCTCAAGTCTCCAAACATCCAGACCCA 873  
QY 661 TCACAAACCTCATCTCCACCAACACACAGCAGTACTGATTAATAAAAA 710  
DB 874 TCACAAACCTCATCTCCACCAACACACAGCAGTACTGATTAATAAAAA 923

RESULT 9  
HRSVGL16 922 bp RNA linear VRL 05-JUN-1997  
LOCUS Human respiratory syncytial virus (MON-7-91) subgroup A, G gene for glycoprotein.  
DEFINITION  
ACCESSION 233429.1 GI:485888  
VERSION G gene: glycoprotein.  
KEYWORDS Human respiratory syncytial virus.  
SOURCE Human respiratory syncytial virus.  
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Dopazo, J.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-1994) Dopazo, J., Centro Nacional de Biotecnologia CSIC, Bioinformatica, Universidad Autonoma, Cantoblanco, Madrid, SPAIN, 28049  
2 (bases 1 to 922)  
AUTHORS Garcia, O., Martin, M., Dopazo, J., Arbizu, J., Fabra, S., Russi, J., Hortel, M., Perez-Brena, P., Martinez, I., Garcia-Barreno, B. and Melero, J.A.  
TITLE Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein  
JOURNAL J. Virol. 68 (9), 5448-5459 (1994)  
MEDLINE 9433057  
PUBMED 8057427

FEATURES  
Source  
Location/Qualifiers  
1..922  
/organism="Human respiratory syncytial virus"  
/strain="subgroup A"  
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"  
/db\_xref="taxon:11250"  
/cell\_line="HeP-2"  
16..912  
/gene="G"  
16..912  
/partial  
/gene="G"  
/codon\_start=1  
/product="glycoprotein"  
/protein\_id="CAA83872.1"  
/db\_xref="GI:485888"  
/db\_xref="SPTREMBL:082068"  
/translation="ISKNDORTAKLEKWTDLNLYLFISSGLYLNKLSIAQITLS  
ILAMISTSLITAIIFLASAHKVTALTATIDQTSQIKNTPTLYLRODOLGISFS  
NLSITSCOTTLASTPGVKNLOPTVKTKNTTGTGPOPSKPTTKROKRPKPKP  
NDHFREVFNFVSCISNNPTMAICKKIPKPKKTKTKTKTKTKTKTKTKTK  
TKRKVEPTTKRKPTTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK  
SOVSTSEHPSPSPPTTRQ"

BASE COUNT 378 a 291 c 97 g 156 t  
Query Match 99.2%; Score 709; DB 14; Length 922;

Best Local Similarity 100.0%; Pred. No. 3,4e-148;  
Matches 709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACACTGCAATCATCAACATGCAACAGCCAGATCAAGAACACA 60  
DB 214 CACAAAGTCACACTAACACTGCAATCATCAACATGCAACAGCCAGATCAAGAACACA 273  
QY 61 ACCCAACATACCTACTCAGATGCTCAGCTTGGAATGAGCTTCTCAATCTGTGAA 120  
DB 274 ACCCAACATACCTACTCAGATGCTCAGCTTGGAATGAGCTTCTCAATCTGTGAA 333  
QY 121 ATTACATCAACACACACCATCACTAGCTTCAACAAACACAGAGATCAAGTCAAACTG 180  
DB 334 ATTACATCAACACACACCATCACTAGCTTCAACAAACACAGAGATCAAGTCAAACTG 393  
QY 181 CAACCCAAACAGTCAGACTTAAACACACACACACACACACACACACACACACAC 240  
DB 394 CAACCCAAACAGTCAGACTTAAACACACACACACACACACACACACACACACAC 453  
QY 241 ACTACAAACACAGCCAAACACACACACACACACACACACACACACACACATATGATTTCACTTGAA 300  
DB 454 ACTACAAACACAGCCAAACACACACACACACACACACACACACACACACATATGATTTCACTTGAA 513  
QY 301 GTGTTTAACCTTTGTAACCTTGAGATATGACAAACATCCACCTGCTGGCTATCTGC 360  
DB 514 GTGTTTAACCTTTGTAACCTTGAGATATGACAAACATCCACCTGCTGGCTATCTGC 573  
QY 361 AAAAGATATCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 420  
DB 574 AAAAGATATCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 633  
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 634 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693  
QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACACCAACCAACCAACCAACCAACCA 540  
DB 694 ACCACCAAGCCACAGAGAGCCCAACCATCAACACCAACCAACCAACCAACCAACCA 753  
QY 541 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 600  
DB 754 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 813  
QY 601 TCACACTCTCCGAGAGCAATCAAGCCTTCTCAAGTCTCCAAACATCCAGACCCA 660  
DB 814 TCACACTCTCCGAGAGCAATCAAGCCTTCTCAAGTCTCCAAACATCCAGACCCA 873  
QY 661 TCACAAACCTCATCTCCACCAACACACAGCAGTACTGATTAATAAAAA 709  
DB 874 TCACAAACCTCATCTCCACCAACACACAGCAGTACTGATTAATAAAAA 922

RESULT 10  
RSHGLYG 917 bp ss-RNA linear VRL 03-AUG-1993  
LOCUS Human respiratory syncytial virus (subgroup A) attachment protein  
DEFINITION (g) mRNA, complete cds.  
ACCESSION M17212.1 GI:333940  
VERSION M17212.1  
KEYWORDS attachment glycoprotein; surface glycoprotein.  
SOURCE Human respiratory syncytial virus (subgroup A, strain Long), cDNA to viral RNA, clones pL63, pL69, and pA.  
ORGANISM Human respiratory syncytial virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 917)  
AUTHORS Johnson, P.R., Springs, M.K., Olmsted, R.A. and Collins, P.L.  
TITLE The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)  
MEDLINE 87289657  
PUBMED 2441388

COMMENT The exact 5' end of long G mRNA was not determined.  
FEATURES  
Location/Qualifiers  
1..917  
/organism="Human respiratory syncytial virus"  
/db\_xref="taxon:11250"  
16..912  
/note="attachment glycoprotein (G)"  
/codon\_start=1  
/protein\_id="AAA47411.1"  
/db\_xref="GI:333941"  
/translation="MSKNDORTAKTLEKTMTHLFISSGLYKLNKSLAQITLS  
ILAMISLSLITAIITFISANHHKYLTLTAIIODTSQIKNTPTLYLQDPGLISFS  
NLSEITSTPTTILASTTPGVKSNLQPTTVKNTKNTTGTQIPSKPTTKQKRNKPKPN  
NDHEEVENFVPCISISNNPTCMALCKRIPNKPKGKTKTKTKPKTKTKDKHPQ  
TTKPKVEPTTKPEPTTNTTKNTLITLLTNNTGNPKLTSQMETFSTSEGNLSP  
SOVSTSEHPSPSPPNTRQ"  
BASE COUNT 375 a 290 c 97 g 154 t 1 others  
ORIGIN

Query Match 97.8%; Score 699.2; DB 14; Length 917;  
Best Local Similarity 99.6%; Pred. No. 5,2e-146;  
Matches 701; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACAAGTACACTAACAAGTGCATCATGCAAGATGCAACAAGCCAGATCAAGAACA 60  
DB 214 CACAAGTACACTAACAAGTGCATCATGCAAGATGCAACAAGCCAGATCAAGAACA 273  
QY 61 ACCCAACATACCTCACTAGATGATCTCAGCTTGAATAGCTTCCATCTGCTGAA 120  
DB 274 ACCCAACATACCTCACTAGATGATCTCAGCTTGAATAGCTTCCATCTGCTGAA 333  
QY 121 ATTACATCAACAACGACCACTACTAGCTTCAACAACAACGAGAGTCAAGTCAACCTG 180  
DB 334 ATTACATCAACAACGACCACTACTAGCTTCAACAACAACGAGAGTCAAGTCAACCTG 393  
QY 181 CAACCCACAACGACCACTAAGACTAAAAACACAACAACCCAAACCAACGAGCAAGCCC 240  
DB 394 CAACCCACAACGACCACTAAGACTAAAAACACAACAACCCAAACCAACGAGCAAGCCC 453  
QY 241 ACTACAAAAACAAGCCAAAAACAACCAACCAACCAATATGATTTTCACTTCGAA 300  
DB 454 ACTACAAAAACAAGCCAAAAACAACCAACCAACCAATATGATTTTCACTTCGAA 513  
QY 301 GTGTTAACTTTGATACCTTGCAGCATATGACAGCAACATTCACACCTGCTATCTGC 360  
DB 514 GTGTTAACTTTGATACCTTGCAGCATATGACAGCAACATTCACACCTGCTATCTGC 573  
QY 361 AAAAGAATACCAAAACAAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 480  
DB 574 AAAAGAATACCAAAACAAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 633  
QY 421 ACCTTCAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 480  
DB 634 ACCTTCAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 693  
QY 481 ACCACCAAGCCCAAGAGAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 540  
DB 694 ACCACCAAGCCCAAGAGAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 753  
QY 541 CTGCTCACCACCAACACCAAGAGAGCCAAACATCAACAGTCAAGTCAAGTCAAGTCAAG 600  
DB 754 CTGCTCACCACCAACACCAAGAGAGCCAAACATCAACAGTCAAGTCAAGTCAAGTCAAG 813  
QY 601 TCAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 660  
DB 814 TCAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 873  
QY 661 TCACAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 704  
DB 874 TCACAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 917

RESULT 11

AR080424  
LOCUS AR080424 894 bp DNA Linear PAT 31-AUG-2000  
DEFINITION Sequence 28 from patent US 5968776.  
ACCESSION AR080424  
VERSION AR080424.1 GI:10007159  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1 (bases 1 to 894)  
AUTHORS Klein,M.H., Du,R.-P. and Ewasysbyn,M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 5968776-A 28 19-OCT-1999;  
FEATURES  
source  
BASE COUNT 364 a 288 c 93 g 149 t  
ORIGIN

Query Match 97.3%; Score 696; DB 6; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.7e-145;  
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTACACTAACAAGTGCATCATGCAAGATGCAACAAGCCAGATCAAGAACA 60  
DB 199 CACAAGTACACTAACAAGTGCATCATGCAAGATGCAACAAGCCAGATCAAGAACA 258  
QY 61 ACCCAACATACCTCACTAGATGATCTCAGCTTGAATAGCTTCCATCTGCTGAA 120  
DB 259 ACCCAACATACCTCACTAGATGATCTCAGCTTGAATAGCTTCCATCTGCTGAA 318  
QY 121 ATTACATCAACAACGACCACTACTAGCTTCAACAACAACGAGAGTCAAGTCAACCTG 180  
DB 319 ATTACATCAACAACGACCACTACTAGCTTCAACAACAACGAGAGTCAAGTCAACCTG 378  
QY 181 CAACCCACAACGACCACTAAGACTAAAAACACAACAACCCAAACCAACGAGCAAGCCC 240  
DB 379 CAACCCACAACGACCACTAAGACTAAAAACACAACAACCCAAACCAACGAGCAAGCCC 438  
QY 241 ACTACAAAAACAAGCCAAAAACAACCAACCAACCAATATGATTTTCACTTCGAA 300  
DB 439 ACTACAAAAACAAGCCAAAAACAACCAACCAACCAATATGATTTTCACTTCGAA 498  
QY 301 GTGTTAACTTTGATACCTTGCAGCATATGACAGCAACATTCACACCTGCTATCTGC 360  
DB 499 GTGTTAACTTTGATACCTTGCAGCATATGACAGCAACATTCACACCTGCTATCTGC 558  
QY 361 AAAAGAATACCAAAACAAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 420  
DB 559 AAAAGAATACCAAAACAAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 618  
QY 421 ACCTTCAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 480  
DB 619 ACCTTCAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGATACC 678  
QY 481 ACCACCAAGCCCAAGAGAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 540  
DB 679 ACCACCAAGCCCAAGAGAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 738  
QY 541 CTGCTCACCACCAACACCAAGAGAGCCAAACATCAACAGTCAAGTCAAGTCAAGTCAAG 600  
DB 739 CTGCTCACCACCAACACCAAGAGAGCCAAACATCAACAGTCAAGTCAAGTCAAGTCAAG 798  
QY 601 TCAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 660  
DB 799 TCAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 858  
QY 661 TCACAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 696  
DB 859 TCACAACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACAATCGAGACCCA 894

RESULT 12	LOCUS	AR092548	894 bp	DNA	Linear	PAT 08-SEP-2000
AR092548	DEFINITION	Sequence 28 from patent US 5998169.				
	ACCESSION	AR092548				
	VERSION	AR092548.1	GI:10019302			
	KEYWORDS					
	SOURCE	Unknown.				
	ORGANISM	Unknown.				
	REFERENCE	Unclassified.				
	AUTHORS	1 (bases 1 to 894)				
	TITLE	Klein,M.H., Du,R.-P. and Ewasysbyn,M.E.				
		Multimeric hybrid gene encoding a chimeric protein which confers				
		protection against parainfluenza virus and respiratory syncytial				
		virus				
		Patent: US 5998169-A 28 07-DEC-1999;				
JOURNAL	FEATURES	Location/Qualifiers				
	SOURCE	1..894				
		/organism="unknown"				
BASE COUNT		364 a 288 c 93 g 149 t				
ORIGIN						
Query Match		97.3%; Score 696; DB 6; Length 894;				
Best local similarity		100.0%; Pred. No. 2,7e-145;				
Matches		696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CACAAAGTCACGTAAACAATCTGAATCTATACAGATGCACACAGCCAGATCAAGACACA	60			
Db	199	CACAAAGTCACGTAAACAATCTGAATCTATACAGATGCACACAGCCAGATCAAGACACA	258			
QY	61	ACCCCAACATACCTCAGTACAGATGTCCTCGTGGATGAGTTCCTCAATCTGTCTGAA	120			
Db	259	ACCCCAACATACCTCAGTACAGATGTCCTCGTGGATGAGTTCCTCAATCTGTCTGAA	318			
QY	121	ATTATATCAACAACACACCATCTACTAGTCTTAACAACAACACAGAGTCAAGTCAACTCG	180			
Db	319	ATTATATCAACAACACACCATCTACTAGTCTTAACAACAACACAGAGTCAAGTCAACTCG	378			
QY	181	CAACCCACACAGTCAAGACTAAAAACACACACACACACACACACACACACACACAC	240			
Db	379	CAACCCACACAGTCAAGACTAAAAACACACACACACACACACACACACACACACAC	438			
QY	241	ACTACAAAACAAGCCAAAACCAACCCAAACCAATATGATTTTCACTTGGAA	300			
Db	439	ACTACAAAACAAGCCAAAACCAACCCAAACCAATATGATTTTCACTTGGAA	498			
QY	301	GGTTTAACTTTGTACCCCGCAGCATATGCACACAATCCAAACCTGGGGCTATCTGC	360			
Db	499	GGTTTAACTTTGTACCCCGCAGCATATGCACACAATCCAAACCTGGGGCTATCTGC	558			
QY	361	AAAGATATCAACAACCAAGGAAAGAAACCCACGACGACGACCTTCAAAAAACCA	420			
Db	559	AAAGATATCAACAACCAAGGAAAGAAACCCACGACGACGACCTTCAAAAAACCA	618			
QY	421	ACCTTCAACACACCAAAAAAGATCTCAAACTCTAAACCATTAACCAAGAAAGTAAACC	480			
Db	619	ACCTTCAACACACCAAAAAAGATCTCAAACTCTAAACCATTAACCAAGAAAGTAAACC	678			
QY	481	ACCACCAAGCCACAGAAAGGCAACCATCAACACACACACCAACCAACCAACATCACTACA	540			
Db	679	ACCACCAAGCCACAGAAAGGCAACCATCAACACACACACCAACCAACCAACATCACTACA	738			
QY	541	CTGGCTACCAACAACACACAGAAATTCAAAATCTACAAGTCAATGAAACCTTCCAC	600			
Db	739	CTGGCTACCAACAACACACAGAAATTCAAAATCTACAAGTCAATGAAACCTTCCAC	798			
QY	601	TCAACCTCTCTCGAAGAGCAATCTAAACCTCTTCAAGTCTCCCAACATCTCGAGACCA	660			
Db	799	TCAACCTCTCTCGAAGAGCAATCTAAACCTCTTCAAGTCTCCCAACATCTCGAGACCA	858			
QY	661	TCACAAACCTCATCTTCACACCAACACACGCGAC	696			

DB	859	TCACAAACCCATCATCTCCACCAACACACACACGACCGC	894
RESULT 13	RS1ICE	8510 bp ss-RNA	linear VRL 29-NOV-2000
LOCUS	Human respiratory syncytial virus nonstructural protein (1C), nonstructural protein (1B), major nucleocapsid (N), phosphoprotein (P), protein (M), 1A (1A), G (G), protein (F) and envelope-associated protein (22K) gene, complete cds		
DEFINITION	M11486 K01445 K02719 K03348 K03349 M11217 M11244 M11487 M11505 M11514 M11631 M12966		
ACCESSION	M11486.1 GI:333925		
VERSION			
KEYWORDS	envelope-associated protein; fusion glycoprotein; major nucleocapsid protein; major surface glycoprotein; matrix protein; nonstructural protein; phosphoprotein.		
SOURCE	Human respiratory syncytial virus.		
ORGANISM	Human respiratory syncytial virus		
REFERENCE	Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.		
AUTHORS	1 (bases 1085 to 2220; 5254 to 5544)		
TITLE	Elango, N. and Venkatesan, S.		
JOURNAL	Amino acid sequence of human respiratory syncytial virus nucleocapsid protein		
MEDLINE	Nucleic Acids Res. 11 (17), 5941-5951 (1983)		
PUBMED	6310521		
REFERENCE	2 (bases 3211 to 4157)		
AUTHORS	Satake, M. and Venkatesan, S.		
TITLE	Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein		
JOURNAL	J. Virol. 50 (1), 92-99 (1984)		
MEDLINE	84138836		
PUBMED	6699948		
REFERENCE	3 (bases 2288 to 3191)		
AUTHORS	Satake, M., Elango, N. and Venkatesan, S.		
TITLE	Sequence analysis of the respiratory syncytial virus phosphoprotein gene		
JOURNAL	J. Virol. 52 (3), 991-994 (1984)		
MEDLINE	85033973		
PUBMED	6548527		
REFERENCE	4 (bases 5602 to 7500)		
AUTHORS	Collins, P.L., Huang, Y.T. and Wertz, G.W.		
TITLE	Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)		
MEDLINE	85088471		
PUBMED	6096849		
REFERENCE	5 (bases 7551 to 8510)		
AUTHORS	Collins, P.L. and Wertz, G.W.		
TITLE	The envelope-associated 22K protein of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polytranscript		
JOURNAL	J. Virol. 54 (1), 65-71 (1985)		
MEDLINE	85135082		
PUBMED	3838351		
REFERENCE	6 (bases 8 to 1050; 7554 to 8506)		
AUTHORS	Elango, N., Satake, M. and Venkatesan, S.		
TITLE	mRNA sequence of three respiratory syncytial virus genes encoding two nonstructural proteins and a 22K structural protein		
JOURNAL	J. Virol. 55 (1), 101-110 (1985)		
MEDLINE	85237684		
PUBMED	4009789		
REFERENCE	7 (bases 5602 to 7423)		
AUTHORS	Elango, N., Satake, M., Colligan, J.E., Norrby, E., Camargo, E. and Venkatesan, S.		
TITLE	Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of FI subunit		
JOURNAL	Nucleic Acids Res. 13 (5), 1559-1574 (1985)		
MEDLINE	85215565		
PUBMED	2987829		
REFERENCE	8 (bases 4630 to 5543)		
AUTHORS	Satake, M., Colligan, J.E., Elango, N., Norrby, E. and Venkatesan, S.		



TITLE Respiratory syncytial virus envelope glycoprotein (G) has a novel structure  
JOURNAL Nucleic Acids Res. 13 (21), 7795-7812 (1985)  
MEDLINE 86067198  
PUBMED 4069997  
REFERENCE 9 (bases 4627 to 5544)  
AUTHORS Wertz,G.W., Collins,P.L., Huang,Y., Gruber,C., Levine,S. and Ball,L.A.  
TITLE Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (12), 4075-4079 (1985)  
MEDLINE 85216636  
PUBMED 3858865  
REFERENCE 10 (bases 4173 to 4571)  
AUTHORS Collins,P.L. and Wertz,G.W.  
TITLE The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript  
JOURNAL Virology 141 (2), 283-291 (1985)  
MEDLINE 86098645  
PUBMED 3879976  
REFERENCE 11 (bases 1 to 528; 552 to 1050)  
AUTHORS Collins,P.L. and Wertz,G.W.  
TITLE Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus  
JOURNAL Virology 143 (2), 442-451 (1985)  
MEDLINE 86045905  
PUBMED 2998021  
REFERENCE 12 (bases 1081 to 2277)  
AUTHORS Collins,P.L., Anderson,K., Janger,S.J. and Wertz,G.W.  
TITLE Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus  
JOURNAL Virology 146 (1), 69-77 (1985)  
MEDLINE 85301974  
PUBMED 3839952  
REFERENCE 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)  
AUTHORS Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camarero,E. and Coelingh,K.V.  
TITLE Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)  
MEDLINE 86259643  
PUBMED 3460060  
COMMENT Reprint for [2] and clean copy sequence for [12],[5] kindly provided by P.Collins, 21-FEB-1986.  
FEATURES  
source location/Qualifiers  
1..8510  
/organism="Human respiratory syncytial virus"  
/strain="A2"  
/db\_xref="taxon:11250"  
/lab\_host="HEp-2 cell"  
1..528  
/gene="1C"  
1..528  
/gene="1C"  
55..474  
/gene="1C"  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAB59850.1"  
/db\_xref="GI:333926"  
/translation="MGSNSLSMIKVRILQNLFDNDEVALIKITCTYDKLIHLTNALANA

VIHITKLNGIVFVHTTSSDICPNNNIVKSNFTMPVLONGYIWEEMELTHCSQPN  
GLDNDNCEIKFSKSLSDSTMTWNNQJSELLGFDLNP"  
171  
/gene="1C"  
/note="c in [11]; t in [6]"  
/replace="c"  
552..1050  
/gene="1B"  
552..1050  
/gene="1B"  
584..958  
/gene="1B"  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAB59851.1"  
/db\_xref="GI:333927"  
/translation="MDTTHNDNTQRLMITDMRPLSELTITSITLRDITTHKTYLIN  
HECIVRKDEQATFPLVNTNEMKLLHKVSTKYKRYTEYNTXGYTPMPDIFTHDGF  
LECGIKPTKHTPIIVKYDLPN"  
747  
/gene="1B"  
/note="a in [11]; g in [6]"  
/replace="g"  
769  
/gene="1B"  
/note="g in [11]; a in [6]"  
/replace="g"  
1081..2277  
/gene="N"  
1081..2277  
/gene="N"  
1096..2271  
/gene="N"  
/codon\_start=1  
/product="major nucleocapsid"  
/protein\_id="AAB59852.1"  
/db\_xref="GI:333928"  
/translation="MALSKYKLNLTDLKDDLLSSSKYT.QRSYQDSIDPNNYDOKHI  
NKLGLLITLEDANHKFTGLIGMLYAMSRIGREDTITLQWYHYKANGVDYTHRO  
DINKREKPFVLTLASLTLEIOTIEIESRSYKMKLKEGEVAPERYHSPGCMIT  
LCIALVITKLAQDRSLTAVIRANNVINKMKRYKGLLPDIANSIPEVEKEKPH  
FIDVFVHFGILOSSTRGCSRVEGITFAGLEKNATGAGOVMLRWCVLAKSVANILGHS  
VOAMEQVVEVEYEAQRLGGAGYVHLNPKASLSLTQFPHFSSVVLGNAGLGLIM  
GEGYRPNQDLDAKAYAEOLKENGVINYSVLDLTAELLEAIKIQLNKDMDEL"  
2289..3196  
/gene="p"  
2289..3196  
/gene="p"  
2302..3027  
/gene="p"  
/codon\_start=1  
/product="phosphoprotein"  
/protein\_id="AAB59853.1"  
/db\_xref="GI:333929"  
/translation="MEKFAPEFHGEDDANNRATFLESIKGFTSPDPKKSIIISVN  
SIDLEVKRESPITSNPIINPTNETDGTAGKNPYORKP.VSFEKDPSPDNFSLY  
KETLETPDNNSESSYSYEIINDQNTDITARLDRIKDELITLGMHTLVVASGPT  
SARGDITDAMIGLREEMIEKIRTEALMTNRLLEAMRLRNESEKMAKOTSDVSLNP  
TSEKLNLLLEGNSDNDLSIEDP"  
Query Match 91.6%; Score 654.8; DB 14; Length 8510;  
Best local similarity 94.8%; Pred. No. 3; 6e-136;  
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 1 CACAAGTCACAGTAAACGTCATCATACAGTCAACCAAGCCAGATCAAGACACA 60  
|||||  
DB 4839 CACAAGTCACACCAACAGTCATCATCAAGATGACCAACCCAGATCAAGACACA 4998  
QY 61 ACCCAACATACCTCAGTCAAGATCTCAGCTTGATGACGCTTCCTCAATCTGTGAA 120  
|||||  
DB 4899 ACCCAACATACCTCAGTCAAGATCTCAGCTTGATGACGCTTCCTCAATCTGTGAA 4958  
QY 121 ATTACATCACAAACCCACCATCTACTAGCTTCAACAACACCGAGATCAAGTCAAACTG 180

```

Db 4959 ATTACATCAAAATACCGACCTACTAGCTTAAACACACCGAGGTCAAGTCAACCTG 5018
Oy 181 CAACCCACAAGTCAAGACTAAAAACACACACACCAACCAACCAAGAGCC 240
Db 5019 CAATCCACAAGTCAAGTCAAGCAAAAAACACACACACTCAACCAACCAAGAGCC 5078
Oy 241 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5079 ACCACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5138
Oy 301 GTGTTAACTTTGTACCTTGACGATATGACAGCAACCACTGCTGGGCTATCTGC 360
Db 5139 GTGTTAACTTTGTACCTTGACGATATGACAGCAACCACTGCTGGGCTATCTGC 5198
Oy 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAAGCTTCAAAAAACCA 420
Db 5199 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAAGCTTCAAAAAACCA 5258
Oy 421 ACCTTCAAGCAACCAAAACCAAGTCAACCTCAACCACTTCAACCAAGGAGTACCC 480
Db 5259 ACCTTCAAGCAACCAAAACCAAGTCAACCTCAACCACTTCAACCAAGGAGTACCC 5318
Oy 481 ACCACCAAGCCCAAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
Db 5319 ACCACCAAGCCCAAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5378
Oy 541 CTGCTACCAACCAACCAACCAAGAAATCCAAATCCAAAGTCAAAATGGAACCTTCCAC 600
Db 5379 CTACTCACTCCCAACCAACCAAGAAATCCAAAGTCAAAATGGAACCTTCCAC 5438
Oy 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660
Db 5439 TCAACTTCTCCGAAGGCAATCCAAAGCCCTTCTCAAGTCTCCACACATCCGAGTACCA 5498
Oy 661 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
Db 5499 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5552

RESULT 14
LOCUS AR093219 15222 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 23 from patent US 5998602.
ACCESSION AR093219
VERSION AR093219.1 GI:10019970
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 15222)
AUTHORS Torrence,P.F., Silverman,R.Hugh., Cirino,N.Mario., Li,G. and Xiao,W.
TITLE Rnaase I activators and antisense oligonucleotides effective to treat RSV infections
JOURNAL Patent: US 5998602-A 23 07-DEC-1999;
FEATURES
source 1..15222
location/Qualifiers
BASE COUNT 5924 a 2704 c 2356 g 4238 t
ORIGIN
Query Match 91.6%; Score 654.8; DB 6; Length 15222;
Best local Similarity 94.8%; Pred. No. 3.4e-136;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Oy 1 CACAAAGTCACTAACAATGCAATCAATCAAGATGCAACCAAGCCAGATCAAGAACACA 60
Db 4886 CACAAAGTCACTAACAATGCAATCAATCAAGATGCAACCAAGCCAGATCAAGAACACA 4945
Oy 61 ACCCAACATACGCTAGATGCTCGAGCTTGAATCAAGCTTCCATCAGTGTGTGA 120
Db 4946 ACCCAACATACGCTAGATGCTCGAGCTTGAATCAAGCTTCCATCAGTGTGTGA 5005

```

```

Oy 121 ATTACATCAAAACCAACCACTATAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 180
Db 5006 ATTACATCAAAACCAACCACTATAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 5065
Oy 181 CAACCCACAAGTCAAGACTAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240
Db 5066 CAATCCACAAGTCAAGTCAAGCAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5125
Oy 241 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5126 ACCACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185
Oy 301 GTGTTAACTTTGTACCTTGACGATATGACAGCAACCACTGCTGGGCTATCTGC 360
Db 5186 GTGTTAACTTTGTACCTTGACGATATGACAGCAACCACTGCTGGGCTATCTGC 5245
Oy 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAAGCTTCAAAAAACCA 420
Db 5246 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAAGCTTCAAAAAACCA 5305
Oy 421 ACCTTCAAGCAACCAAAACCAAGTCAACCTCAACCACTTCAACCAAGGAGTACCC 480
Db 5306 ACCTTCAAGCAACCAAAACCAAGTCAACCTCAACCACTTCAACCAAGGAGTACCC 5365
Oy 481 ACCACCAAGCCCAAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
Db 5366 ACCACCAAGCCCAAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5425
Oy 541 CTGCTACCAACCAACCAACCAAGAAATCCAAATCCAAAGTCAAAATGGAACCTTCCAC 600
Db 5426 CTACTCACTCCCAACCAACCAAGAAATCCAAAGTCAAAATGGAACCTTCCAC 5485
Oy 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660
Db 5486 TCAACTTCTCCGAAGGCAATCCAAAGCCCTTCTCAAGTCTCCACACATCCGAGTACCA 5545
Oy 661 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
Db 5546 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5599

RESULT 15
LOCUS HRU50362 15222 bp RNA linear VRL 20-NOV-1997
DEFINITION Human respiratory syncytial virus, mutant cp-RSV, complete genome.
ACCESSION U50362
VERSION U50362.1 GI:2627296
KEYWORDS
SOURCE Human respiratory syncytial virus.
ORGANISM Human respiratory syncytial virus.
REFERENCE 1 (bases 1 to 15222)
AUTHORS Paranyxoviridae; Pneumovirinae; Pneumovirus.
TITLE 1 (bases 1 to 15222)
JOURNAL Collins,P.L.
Virus contians mutations in the F and L genes
Virology 208 (2), 478-484 (1995)
MEDLINE 9526253
PUBMED 7747420
REFERENCE 2 (bases 1 to 15222)
AUTHORS Crowe,J.E. Jr., Firestone,C.Y., Whitehead,S.S., Collins,P.L. and Murphy,B.R.
Acquisition of the ts phenotype by a chemically mutagenized cold-passaged human respiratory syncytial virus vaccine candidate results from the acquisition of a single mutation in the polymerase (L) gene
Virus Genes 13 (3), 269-273 (1996)
MEDLINE 97187925
PUBMED 9035372
REFERENCE 3 (bases 1 to 15222)
AUTHORS Whitehead,S.S.

```

TITLE Direct Submission  
JOURNAL Submitted (29-FEB-1996) RVS, LID, NIAID, Bldg 7, Rm 118, 7 Center  
Drive, MSC 0720, Bethesda, MD 20892, USA  
FEATURES  
source 1.15222  
/organism="Human respiratory syncytial virus"  
/strain="A2"  
/db\_xref="taxon:11250"  
/note="mutant cp-RSV"  
45..576  
/gene="NS1 (1C)"  
45..576  
/gene="NS1 (1C)"  
/product="nonstructural protein NS1"  
99..518  
/gene="NS1 (1C)"  
/codon\_start=1  
/product="nonstructural protein NS1"  
/protein\_id="AAB86656.1"  
/db\_xref="GI:2627297"  
/translation="MGSNSLSMKVRLQNLFDNDEVALIKITCYNDKLIHLTNALAKA  
VIHTIKINGIVFVHVTSSDLCPPNNIVKSNFTPYLQNGYIMEMMELTHCSQPN  
GLDNCCEIKFKSKLSDSTMVYMNQISELGLFDLNP"  
596..1098  
/gene="NS2 (1B)"  
596..1098  
/gene="NS2 (1B)"  
/product="nonstructural protein NS2"  
628..1002  
/gene="NS2 (1B)"  
/codon\_start=1  
/product="nonstructural protein NS2"  
/protein\_id="AAB86657.1"  
/db\_xref="GI:2627298"  
/translation="MDTTHNDNTPQRLMTMPRLSFTITSLTRDITLTKFTYLLN  
HECIVRLDEQATPTPLVNEEMKLHKVSGTKIKTYEYNTKIGTFPMPPIFINHDF  
LECIKPKTKHTPIYKYDLNP"  
1123..2327  
/gene="N"  
1125..2327  
/gene="N"  
/product="major nucleocapsid"  
1140..2315  
/gene="N"  
/codon\_start=1  
/product="major nucleocapsid"  
/protein\_id="AAB86658.1"  
/db\_xref="GI:2627299"  
/translation="MALSQVLTNLTNKLQDLSSSKYITQSTGDSIDTPNYVQKHI  
NKLGMILITDANKRTFTGLIGMLTAMSLGREDTIKLRDAGVHVANGVDVTHRO  
DINCKEMKEFVLTLASLTETQINIEIESRSYKMKKEMGEVAPVRRHSDPCGMII  
LCIALVLTKLAAQDRSGLTAVIRRANVLRKEMKRYKGLPKDIANSFYVFKHPH  
FIDVVFHGIAQSDRSGSRVGEIGFAGLFNNAVGAGQMLRMGVLAKSIKIMIGHNS  
VOAEGEYQVEVEYFAOKLGEAGFYHILNNKASLISLTQPPHSSVYLGAALGIM  
GEYRGTPRNQDLYAAKAYAQOLKENCYINISVLDLTAELLEAIKROLNPRKNDVEL"  
2329..3242  
/gene="P"  
2329..3242  
/gene="P"  
/product="phosphoprotein"  
2346..3071  
/gene="P"  
/codon\_start=1  
/product="phosphoprotein"  
/protein\_id="AAB86659.1"  
/db\_xref="GI:2627300"  
/translation="MEKFAPEFHGEDANNRATKFLSEIKGKFTSPKDKKSDSIYSN  
SIDLEVTRESPTNSYIINPTNETDPTAGKPKYQKRLPVSEFEDTPSPDNPSKLY  
KETLETPDNNEESYSYEBINDOTNDNTTARDRIDEKUSELIGMLHTLVASAGPT  
SARDGITDANGLKEEMIEKIRLTALNTDRLFAMARLRNEESEKMAKDTSDVSLNP  
TSEKLNLILEGSDNDLSLEDF"  
3252..4209  
/gene="M"  
3252..4209  
/gene="M"

3252..4209  
/gene="M"  
/product="Matrix protein M"  
3261..4031  
/gene="M"  
/codon\_start=1  
/product="Matrix protein M"  
/protein\_id="AAB86660.1"  
/db\_xref="GI:2627301"  
/translation="METVYVKLHEGSTYTAAYQNVLEKDDPASTITWYPMQSSNP  
ADLILKELANVILVQISTPGSLRWVAINSRSAVLACMPKFTICANVSDERKLI  
AYDVTPCEIKACSLTCLSKSNMLITVQDLTMETIHDIILCEPENIVTSKVIYI  
PTYRSISVRKNDLNTLENTTTEFKAIINAKIIPYSGLIYITVTDNKGAKYIKP  
OSQPIVDLGALEKESTIYYTTNNKHATPAIKPMED"  
3970..4197  
/gene="M"  
/note="putative alternative translation of M gene"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAB86661.1"  
/db\_xref="GI:2627302"  
/translation="MLPQIGSTQLHDIQSPMKINLPLHQCYNYSKLSLYLHFTIT  
ITNTLWPNQSKTYIKSQIIPSHCLSDVLK"  
4219..4628  
/gene="SH (1A)"  
4219..4628  
/gene="SH (1A)"  
/product="small hydrophobic protein"  
4303..4497  
/gene="SH (1A)"  
/codon\_start=1  
/product="small hydrophobic protein"  
/protein\_id="AAB86662.1"  
/db\_xref="GI:2627303"  
/translation="MENVSTITIEPSSKFWPYFLIMHTITISLLIISIMAILNKL  
CEYNVFNKTFELPARAVNT"  
4673..5595  
/gene="G"  
4673..5595  
/gene="G"  
/product="attachment G glycoprotein"  
4688..5584  
/gene="G"  
/codon\_start=1  
/product="attachment G glycoprotein"  
/protein\_id="AAB86663.1"  
/db\_xref="GI:2627304"  
/translation="MSKNKQDQRAKTLERTWDTLNLHLFTSSCYLKLNSVAQITLS  
ILMILISTSLIAAIIFFIASANHKVPTPAI IODATSOIKNTPTVLTQNPOLGISPS  
NPSEILSQITITLASTTPGVKSTLQSTYVTKNTTQTPQSKPTKOROKNPSPKN  
NDHFEEVFNVPSCISNNPTCAICRKRIPNKKPGKTTTKPTKPTLTKTKKDPQO  
TTSKEVPTPKPREPTINTKNTITTLTISMTGNPELTQMETFHTSSGQNSP  
SOVSTTSSEYSSQSSPPNTPRO"  
5648..7550  
/gene="F"  
5648..7550  
/gene="F"  
/product="fusion F glycoprotein precursor"  
5661..7385  
/gene="F"  
/codon\_start=1  
/product="fusion F glycoprotein precursor"  
/protein\_id="AAB86664.1"  
/db\_xref="GI:2627305"  
/translation="MELLILKANAITTILTAVFCFASGQNTIEEPYOSTGSAVSGY  
LSLRKQWTSVTTTLELSINKEKCKGTAQKAKLLKQELDKRKNVTEIQLIMOSTPA  
TNKARRELPRTMNTLNAKKTIVTLNKKRRKREGLIGSATAVNSAVSKVLAH  
EGNRKIKSALLSTNAVSLNSGVSLNSKVLDKNTYIDKQLPLPVNQSQISNIA  
TVIEFOQKNRRLIEITREFSVNAGVTPPVSTYMLNSELISINDMPTINDOKILMSN  
MNOIVQOQSYISINIKIEBVLAAYVOLPLYGVIDPCKMLHTSPICPTNTRKGSNCL  
TRDRGMVQDNAGSVSFPPQAECKVQSNRVCMDMNSLTLPSVNLGNVDIENKDY  
CKLMTSKTDVSSVLTSLGAIYSCVTKCTJASNNKRGILKTFPSCDYSNKGVDYV  
SVGNLTLYYVKQBGKSLYVGEPIINFTDPLVFPSEDFPASTISOVNEKINQSLATIRK

```
SDLLHNYNACKSTINIMITIIIVIIILLSLAVGLLLYCKARSPVTLRSKQSLG
INNTAFSN"
7597..8557
/gene="M2 (22K)"
7597..8557
/gene="M2 (22K)"
/product="matrix protein M2"
7606..8190
/gene="M2 (22K)"
/codon_start=1
/product="matrix protein M2"
/protein_id="AAB8665.1"
/db_xref="GI:2627306"
/translation="MSRRNPCKFEIRGHCLNGKRHPSHNYFEMPYPHALLYRONFMLN
RLIKSMKSIDTLEISGAALDRREYALGVLEISYIGSINNTIKOSACVMSKL
LTELNSDDIKLRONEELNSPKIRYNTVISYIESNRKNKQTHLLKRLPADVLAKT
IKNTLDIHKSTINNPKESIVSDTNDAKNDDT"
8159..8431
/gene="M2 (22K)"
/note="putative alternative translation of M2 gene"
/codon_start=1
/product="unknown"
/protein_id="AAB8666.1"
/db_xref="GI:2627307"
/translation="VTMPKIMILPDKYPGSIYSILISRCRVIMYNOKNTLYFNQNNP

CDS

Query Match      91.6%; Score 654.8; DB 14; Length 15222;
Best Local Similarity 94.8%; Pred. No. 3.4e-136;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY      1  CACAAGTCACACTAATGCAATCATCAAGATGCAACAAGCCAGATCAAGAACA 60
DB      4886 CACAAGTCACACCACTGCAATCATCAAGATGCAACAAGCCAGATCAAGAACA 4945
QY      61  ACCCCACATACCTCAGATCTCTGAGATCTGGAATCAGCTTCTCAATCTGTCTGAA 120
DB      4946 ACCCCACATACCTCAGATCTCTGAGATCTGGAATCAGCTTCTCAATCTGTCTGAA 5005
QY      121 ATTACATCAAAACCAAGCACTAGTTCACACAACCAAGAGTCAAGTCAAACTG 180
DB      5006 ATTACATCAAAATCCACCATCTAGTTCACACAACCAAGAGTCAAGTCAAACTG 5065
QY      181 CAACCCACAACAGTCAAGACTAAACACACAACCAACCAACCAACCAAGCCG 240
DB      5066 CAATCCACAAGTCAAGACCAAAACACACAACCAACCAACCAACCAAGCCG 5125
QY      241 ACTACAACCAAGCCCAAAACCAACCAACCAACCAATATGATTTTCACTTGA 300
DB      5126 ACCACAACCAAGCCCAAAACCAACCAACCAACCAATATGATTTTCACTTGA 5185
QY      301 GTGTTAACTTGTACCTGACATATGACACAATCAACCTGCTGGCTATCTGC 360
DB      5186 GTGTTCACTTGTACCTGACATATGACACAATCAACCTGCTGGCTATCTGC 5245
QY      361 AAAAGAAATCAAAAAACCAAGAAACCAACCAAGCTTACAAAAAACCA 420
DB      5246 AAAAGAAATCAAAAAACCAAGAAACCAACCAAGCTTACAAAAAACCA 5305
QY      421 ACCTTAAGACCAAAAAAGATCTCAAACTCAAACTAAACCAAGAAAGTACCC 480
DB      5306 ACCTTAAGACCAAAAAAGATCTCAAACTCAAACTAAACCAAGAAAGTACCC 5365
QY      481 ACCACAAGCCACAGAAGCCACATCAACACCAACCAACCAACATCAACACTACA 540
DB      5366 ACCACAAGCCACAGAAGCCACATCAACACCAACCAACCAACATCAACACTACA 5425
QY      541 CTGCTACCAACAACACACAGAAATCCAAAATCACAAGTCAATGGAACCTTCAC 600
DB      5426 CTACTACCTCCACACACAGAAATCCAAAATCACAAGTCAATGGAACCTTCAC 5485
QY      601 TCAACCTCCGAGAGCAATTAAGCCCTTCAAGTCTCCACAACATCCGACACCA 660
DB      5486 TCAACTTCTCCGAGAGCAATTAAGCCCTTCAAGTCTCTACACATCCGAGTACCA 5545
```

```
QY      661 TCACACCTTCATCTCCACCAACACACAGCCAGTAGTTATTAACAAAAA 714
DB      5546 TCACACCTTCATCTCCACCAACACACAGCCAGTAGTTATTAACAAAAA 5599
```

Search completed: May 1, 2003, 11:59:54  
Job time : 1461.68 secs

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:53:37 ; Search time 139.502 Seconds  
(without alignments)  
11542.374 Million cell updates/sec

Title: US-09-462-816-3

Sequence: 1 cacaagtcaactaac.....gtagttatttaaaaaaaaaa 715

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_101002:★

1	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1960.DAT.*
2	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1961.DAT.*
3	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1962.DAT.*
4	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1963.DAT.*
5	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1964.DAT.*
6	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1965.DAT.*
7	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1966.DAT.*
8	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1967.DAT.*
9	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1968.DAT.*
10	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1969.DAT.*
11	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1970.DAT.*
12	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1991.DAT.*
13	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1992.DAT.*
14	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1993.DAT.*
15	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1994.DAT.*
16	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1995.DAT.*
17	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1996.DAT.*
18	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1997.DAT.*
19	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1998.DAT.*
20	/SID5/gcgdata/geneseq/geneseqn_emb1/NA1999.DAT.*
21	/SID5/gcgdata/geneseq/geneseqn_emb1/NA2000.DAT.*
22	/SID5/gcgdata/geneseq/geneseqn_emb1/NA2001.DAT.*
23	/SID5/gcgdata/geneseq/geneseqn_emb1/NA2002.DAT.*
24	/SID5/gcgdata/geneseq/geneseqn_emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	715	100.0	715	20	AA08422	G protein gene first
2	715	100.0	920	14	AA045686	Respiratory syncyt
3	715	100.0	920	0	AA08421	Membrane bound G f
4	659	92.2	935	8	AA070784	Sequence encoding
5	659	92.2	935	19	AAV18736	HRV glycoprotein
6	657.4	91.9	935	13	AA026623	HRV glycoprotein
7	654.8	91.6	15210	20	AA059703	Polynucleotide seq
8	654.8	91.6	15222	18	AA078440	Human respiratory
9	654.8	91.6	15223	18	AA065430	Respiratory syncyt

10	654.8	91.6	15223	19	AAV17553	Respiratory syncyt
11	654.8	91.6	15223	21	AAAH8743	Respiratory syncyt
12	649.6	90.9	918	22	AAC88494	Human RSV G-prote
13	646.2	90.4	897	24	AAAS20145	Respiratory syncyt
14	445.8	62.3	696	19	AAAB8298	Respiratory syncyt
15	300.4	42.0	15225	21	AAAV17552	Respiratory syncyt
16	300.4	42.0	15225	21	AAAH8744	Respiratory syncyt
17	300.4	42.0	15225	24	AAD55520	Respiratory syncyt
18	295.6	41.3	15229	19	AAV18276	Human Respiratory
19	295.6	41.3	15229	19	AAV18276	RSV isolate H1537
20	295.6	41.3	15229	20	AAAZ22910	Nucleotide sequenc
21	273.6	38.3	15219	20	AAAX35268	DNA encoding the I
22	273.6	38.3	15219	19	AAAV18275	RSV encoding the I
23	273.6	38.3	15218	20	AAAZ22909	RNA isolate 2B wit
24	273.6	38.3	15218	20	AAAZ22914	Nucleotide sequenc
25	273.6	38.3	15218	20	AAAX35267	Nucleotide sequenc
26	273.6	38.3	15219	19	AAAV18277	DNA encoding the I
27	273.6	38.3	15219	19	AAAV18278	RSV vaccine 2B33F
28	273.6	38.3	15219	19	AAAV18279	RSV vaccine 2B30L
29	273.6	38.3	15219	19	AAAV18280	RSV revertant 2B33
30	273.6	38.3	15219	20	AAAZ22911	RSV revertant 2B20
31	273.6	38.3	15219	20	AAAZ22912	Nucleotide sequenc
32	273.6	38.3	15219	20	AAAZ22913	Nucleotide sequenc
33	273.6	38.3	15219	20	AAAX35269	Nucleotide sequenc
34	273.6	38.3	15219	20	AAAX35270	DNA encoding the I
35	273.6	38.3	15219	20	AAAX35271	DNA encoding the I
36	212.8	29.8	1050	17	AATP31647	DNA encoding the L
37	212.8	29.8	1050	17	AATP31647	Encodes Streptococ
38	212.8	29.8	1356	22	AAF84711	Nucleotide sequenc
39	210.8	29.5	303	16	AAAF01533	Nucleotide sequenc
40	210.8	29.5	303	16	AAAI03466	RSV subgroup A clat
41	210.8	29.5	303	17	AATG27073	RSV subgp. A prote
42	210.8	29.5	303	17	AATG27073	RSV sub-group A wit
43	210.8	29.5	303	20	AAAT31649	Respiratory Syncyt
44	210.8	29.5	303	20	AAZ30478	RSV G protein anti
45	210.8	29.5	303	21	AAAH5862	DNA encoding a G2N
				22	AAAH78459	Nucleotide sequenc

## ALIGNMENTS

XX	RESULT 1
XX	AAx08422
ID	AAx08422 standard; cDNA; 715 BP.
XX	
AC	AAx08422;
XX	
DJ	28-JUN-1999 (first entry)
DE	
DE	G protein gene fragment of respiratory syncytial virus.
XX	
KW	G protein: respiratory syncytial virus; RSV; recombinant vector
KW	vaccine; immune response; immunogenicity; tPA; antibody;
XX	tissue plasminogen activator; ss.
OS	
XX	Respiratory syncytial virus (RSV).
Key	Location/Qualifiers
FH	1..702
FT	/tag= a
FT	/product= "Secreted G protein"
PN	
XX	W09904010-A1.
XX	
XD	28-JAN-1999.
PF	
PE	16-JUL-1998; 98WO-CA00697.
PR	
XX	18-JUL-1997; 97US-0896442.
PA	(CONN-) CONNAUGHT LAB LTD.
XX	
XI	Klein MH, Li X, Sambhara S;

```

xx WP1: 1999-132254/11.
DR P-PSDB: AAW96314.
xx
xx Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis
xx
xx Claim 8; Figure 3; 67pp; English.
xx
xx The respiratory syncytial virus (RSV) G protein can be used in
xx vaccines by inserting the G protein gene into a non-replicating
xx vector. The G protein is placed under the control of alternative
xx signal and expression sequences, for example the chimeric G protein
xx produced may also comprise the signal peptide of tissue plasminogen
xx activator (tPA). The recombinant vector may also comprise sequences
xx upstream of the G protein gene which enhance the G proteins
xx immunoprotective ability. The resulting immunogenic composition will
xx generate antibodies directed against the RSV G protein when
xx administered to a host organism. The composition is useful as a
xx vaccine to immunise against RSV-associated disease, particularly
xx resulting in a balanced Th1/Th2 immune response and for raising Ab,
xx by usual immunisation and cell fusion methods.
xx
SQ Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other:

Query Match      100.0%; Score 715; DB 20; Length 715;
Best Local Similarity 100.0%; Pred. No. 3.2e-162;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 60
DB 1 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 60
OY 61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 120
DB 61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 120
OY 121 ATTACATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180
DB 121 ATTACATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180
OY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
DB 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
OY 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 300
OY 301 GTGTTTAACCTTGTACCTGACGATATGACAGCAACATCAACCTGCTGCT 360
DB 301 GTGTTTAACCTTGTACCTGACGATATGACAGCAACATCAACCTGCTGCT 360
OY 361 AAAAGAATACCAAAACCAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 420
DB 361 AAAAGAATACCAAAACCAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 420
OY 421 ACCTTAAGACCAACCAAAAGATCTCAAAAGCTCAAAAGCTCAAAAGGAG 480
DB 421 ACCTTAAGACCAACCAAAAGATCTCAAAAGCTCAAAAGCTCAAAAGGAG 480
OY 481 ACCACCAAGCCACAGAGAGCCCAACATCAACCAACCAACCAACCAACCA 540
DB 481 ACCACCAAGCCACAGAGAGCCCAACATCAACCAACCAACCAACCAACCA 540
OY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
OY 601 TCACACTCCGCGAAGGCAATCAAGCCCTTCAAGTCCCAACATCCGAGCA 660
DB 601 TCACACTCCGCGAAGGCAATCAAGCCCTTCAAGTCCCAACATCCGAGCA 660

```

```

DB 601 TCACACTCCGCGAAGGCAATCAAGCCCTTCAAGTCCCAACATCCGAGCA 660
OY 661 TCACACTCCGCGAAGGCAATCAAGCCCTTCAAGTCCCAACATCCGAGCA 715
DB 661 TCACACTCCGCGAAGGCAATCAAGCCCTTCAAGTCCCAACATCCGAGCA 715

RESULT 2
AAQ45686
ID AAQ45686 standard; DNA; 920 BP.
XX
XX AAQ45686;
AC 13-JAN-1994 (first entry)
XX
XX 13-JAN-1994 (first entry)
DE Respiratory syncytial virus (RSV) G gene.
XX
XX PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine;
KW ds.
XX
XX Respiratory syncytial virus.
OS
XX
XX Key Location/Qualifiers
FH CDS 8..901
FT /*tag= a
FT /product= RSV G protein
FT misc_feature 89..176
FT /*tag= b
FT /label= Transmembrane anchor domain
XX
XX W09314207-A.
XX
XX 22-JUL-1993.
PD
XX
XX 05-JAN-1993; 93MO-CA00001.
PE
XX
XX 06-JAN-1992; 92GB-0000117.
PR
XX
XX (CONN-) CONNADGHT LAB LTD.
PA
XX
XX Ewasysyn ME, Klein MH;
PI
XX
XX WP1: 1993-243222/30.
DR P-PSDB: AAR39286.
DR
XX
XX Multimeric hybrid genes and their chimeric proteins - are
PT vaccines against multiple pathogenic infections e.g.
PT para-influenza virus and respiratory syncytial virus
XX
XX Claim 11; Figure 7A-7D; 80pp; English.
XX
XX A novel multimeric hybrid gene is used as a vaccine. The gene
XX consists of two gene sequences which are linked and encode antigenic
XX regions, these two sequences being derived from two different
XX pathogens (parainfluenza virus (PIV) and respiratory syncytial virus
XX (RSV)). The gene sequences that are particularly used are those
XX which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and
XX G proteins (AAQ45685, AAQ45686).
XX
SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other:

Query Match      100.0%; Score 715; DB 14; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.4e-162;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 60
DB 1 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 60
OY 206 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 265
DB 206 CACAAAGTCACACTAACAAGTCATACAGATGCAACAGCCAGATCAAGACACA 265
OY 61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 120
DB 61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 120
OY 266 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 325
DB 266 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 325

```

```
QY 121 ATTACATCAAAACCAACCATCTAGCTTCACACAGCAGAGTCAAACTG 180
DB 326 ATTACATCAAAACCAACCATCTAGCTTCACACAGCAGAGTCAAACTG 385
QY 181 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAAGCC 240
DB 386 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAAGCC 445
QY 241 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTAACTTTGACCTGACATATGACGCAACCAATCCAACTGCTGGCTATGTC 360
DB 506 GTGTTAACTTTGACCTGACATATGACGCAACCAATCCAACTGCTGGCTATGTC 565
QY 361 AAAAATATTCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAATATTCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCACCAAGCCCAAGAGCCCAACATCAACACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAGCCCAACATCAACACCAACCAACCAACCAACCA 745
QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805
QY 601 TCAACCTCCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 660
DB 806 TCAACCTCCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 865
QY 661 TCACAAACCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACAAACCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCA 920
```

RESULT 3  
AA08421

ID AA08421 standard; cDNA: 920 BP.

AC AAX08421;

DT 28-JUN-1999 (first entry)

DE Membrane bound G protein gene of respiratory syncytial virus.

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

KW tissue plasminogen activator; ss.

OS Respiratory syncytial virus (RSV).

FT Key Location/Qualifiers

FT CDS 8..904

FT /product= "Membrane bound G protein"

FT W09904010-A1.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-CA00697.

PR 18-JUL-1997; 97US-0896442.

PA (CONN-) CONNBUGHT LAB LTD.

XX Klein MH, Li X, Sambhara S;

```
XX WPI: 1999-132254/11.
DR P-PSDB; AAW6313.
XX Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis
PS Claim 3; Figure 2; 67pp; English.
XX The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternative
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods.
XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;
```

Query Match 100.0%; Score 715; DB 20; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3,4e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CACAAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 60
DB 206 CACAAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 265
QY 61 ACCCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 120
DB 266 ACCCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 325
QY 121 ATTACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 180
DB 326 ATTACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 385
QY 181 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAACCA 240
DB 386 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAACCA 445
QY 241 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTAACTTTGACCTGACATATGACGCAACCAATCCAACTGCTGGCTATGTC 360
DB 506 GTGTTAACTTTGACCTGACATATGACGCAACCAATCCAACTGCTGGCTATGTC 565
QY 361 AAAAATATTCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAATATTCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCACCAAGCCCAAGAGCCCAACATCAACACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAGCCCAACATCAACACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805
QY 601 TCAACCTCCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 660
```





PT Production of human respiratory syncytial virus glyco-protein F or G  
 PT - by culturing eukaryotic host cells transfected with corresponding  
 PT DNA

XX Example 1; Columns 27-28; 17pp; English.

XX The present sequence was used in the development of a novel method  
 CC for the production of human respiratory syncytial virus (HRSV)  
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises  
 CC culturing eukaryotic host cells transfected with an isolated DNA  
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare  
 CC vaccines against HRSV.

XX Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 92.2%; Score 659; DB 19; Length 935;  
 Best Local Similarity 95.1%; Pred. No. 9,5e-149;  
 Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 CACAAGTACACTAACAAGTCAATCAAGATGCAAGCCAGATCAAGAACACA 60  
 DB 214 CACAAGTCAACAAGTCAATCAAGATGCAAGCCAGATCAAGAACACA 273  
 QY 61 ACCCCAACTACCTCAGATGATCCTGAGTGGATGAGTCTTCCATCTGTGAA 120  
 DB 274 ACCCCAACTACCTCAGATGATCCTGAGTGGATGAGTCTTCCATCTGTGAA 333  
 QY 121 ATTACATCAACAAGCCAGATGATGATGATGATGATGATGATGATGATGAT 180  
 DB 334 ATTACATCAACAAGCCAGATGATGATGATGATGATGATGATGATGATGAT 393  
 QY 181 CAACCCACACAGTCAAGACTAAACACACACACACACACACACACACACAC 240  
 DB 394 CAATCCACACAGTCAAGACTAAACACACACACACACACACACACACACAC 453  
 QY 241 ACTACAAACACAGCCAAACACACACACACACACACACACACACACACAC 300  
 DB 454 ACCACAAACACAGCCAAACACACACACACACACACACACACACACACAC 513  
 QY 301 GTGTTAACTTTGTACCTGACGATATGACGACCAATTCACCTGCTGGCTATTC 360  
 DB 514 GTGTTAACTTTGTACCTGACGATATGACGACCAATTCACCTGCTGGCTATTC 573  
 QY 361 AAAAAGATACCAAAACCAAGGAAAGAAACACACACACACACACACACAC 420  
 DB 574 AAAAAGATACCAAAACCAAGGAAAGAAACACACACACACACACACACAC 633  
 QY 421 ACCTTCAGACACCAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAAC 480  
 DB 634 ACCTTCAGACACCAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAAC 693  
 QY 481 ACCACCAAGCCCAAGAGGACCAACATCAACACACACACACACACACAC 540  
 DB 694 ACCACCAAGCCCAAGAGGACCAACATCAACACACACACACACACACAC 753  
 QY 541 CTGCTCAGACCAACACACAGGAAATCCAAATCTCAAGTCTCAAGATCTCG 600  
 DB 754 CTGCTCAGACCAACACAGGAAATCCAAATCTCAAGTCTCAAGATCTCG 813  
 QY 601 TCAACCTCTCCGAAGGCAATCTTAAGCCTTCTCAAGTCTCAAGATCTCG 660  
 DB 814 TCAACCTCTCCGAAGGCAATCTTAAGCCTTCTCAAGTCTCAAGATCTCG 873  
 QY 661 TCACAACTCTCTCCGAAGGCAATCTTAAGCCTTCTCAAGTCTCAAGATCT 715  
 DB 874 TCACAACTCTCTCCGAAGGCAATCTTAAGCCTTCTCAAGTCTCAAGATCT 928

RESULT 6  
 AAQ29623  
 ID AAQ29623 standard; DNA; 935 BP.  
 XX  
 XX AC AAQ29623;

XX 03-MAR-1993 (first entry)  
 DT  
 XX HSRV glycoprotein G (gpG).  
 DE  
 XX Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;  
 KW major capsid protein; N; ss.  
 XX Human respiratory syncytial virus strain A2.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 16..912  
 FT CDS /tag= a  
 FT /label= G\_protein  
 FT misc\_feature 16..22  
 FT /tag= b  
 FT /note= "oligonucleotide used to probe for full length cDNA"  
 FT misc\_feature 898..912  
 FT /tag= c  
 FT /note= "oligonucleotide used to specifically prime the reverse transcription reaction for making the first strand of the cDNA"  
 FT US5149650-A.  
 FT PD 22-SEP-1992.  
 FE 14-JAN-1986; 86US-0818740.  
 FE 14-JAN-1986; 86US-0818740.  
 PR 13-JUL-1988; 88US-0218737.  
 PR (UNCL) UNIV NORTH CAROLINA.  
 PA  
 XX Collins PL, Wertz GW;  
 P1 WPI: 1992-340247/41.  
 DR P-PSDB; AAR25302.  
 DR  
 XX  
 XX Vaccines for human respiratory virus - include structural genes  
 PT coding for native structural viral proteins and immunogenic  
 PT fragments  
 PT  
 PS  
 XX Disclosure; Page 18; 21pp; English.  
 CC The sequences of mRNA encoding HRSV structural proteins are given in  
 CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid  
 CC protein N. The sequences and encoded proteins are useful for  
 CC preparing vaccines against HRSV. The vaccines can be used to confer  
 CC immunity against respiratory tract infections on human subjects.  
 CC  
 XX  
 SO Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;  
 Query Match 91.9%; Score 657.4; DB 13; Length 935;  
 Best Local Similarity 95.0%; Pred. No. 2.3e-148;  
 Matches 679; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 CACAAGTACACTAACAAGTCAATCAAGATGCAAGCCAGATCAAGAACACA 60  
 DB 214 CACAAGTCAACAAGTCAATCAAGATGCAAGCCAGATCAAGAACACA 273  
 QY 61 ACCCCAACTACCTCAGATGATCCTGAGTGGATGAGTCTTCCATCTGTGAA 120  
 DB 274 ACCCCAACTACCTCAGATGATCCTGAGTGGATGAGTCTTCCATCTGTGAA 333  
 QY 121 ATTACATCAACAAGCCAGATGATGATGATGATGATGATGATGATGATGAT 180  
 DB 334 ATTACATCAACAAGCCAGATGATGATGATGATGATGATGATGATGATGAT 393  
 QY 181 CAACCCACACAGTCAAGACTAAACACACACACACACACACACACACACAC 240  
 DB 394 CAATCCACACAGTCAAGACTAAACACACACACACACACACACACACACAC 453



KM Antisense oligonucleotide; respiratory syncytial virus;  
 KM RSV; treatment; infection; inhibition; strain A2; ss.  
 OS Human respiratory syncytial virus.  
 PN W09729757-A1.  
 XX  
 XX 21-AUG-1997.  
 PD  
 XX 14-FEB-1997; 97WO-US02531.  
 PF  
 XX 15-FEB-1996; 96US-0011725.  
 PR  
 XX (CLEV-) CLEVELAND CLINIC FOUND.  
 PA (USSH ) US NAT INST OF HEALTH.  
 XX  
 PI Clitino NM, Li G, Silverman RH, Torrence PF, Xiao W;  
 DR WPI: 1997-424748/39.  
 XX  
 XX Polynucleotide containing sequence anti-sense to region of RSV -  
 PT connected via a linker to an activator of RNaseL, used to treat RSV  
 PT infections  
 XX  
 XX Disclousure; Pages 47-51; 89pp; English.  
 XX  
 XX The present sequence was used in the preparation of a novel  
 CC polynucleotide, comprising an antisense oligonucleotide, with a  
 CC hydroxy group at one end, that is complementary to 15-20 bases of  
 CC the anti-genomic RNA strand of a respiratory syncytial virus (RSV),  
 CC a linker attached to the OH-end of the antisense oligonucleotide and  
 CC an oligonucleotide activator of RNaseL attached to the linker. The  
 CC polynucleotide can be used to treat RSV infections, which can be  
 CC be treated by administration of the antisense oligonucleotide, so as  
 CC to form a complex with activated RNase L in vivo. The  
 CC polynucleotide can be transported across the cell membranes without  
 CC carriers or permeability agents, and once introduced destroys  
 CC antisense target RNA. It also inhibits RSV infection in vitro in a  
 CC superior manner to the conventional drug, ribavirin.  
 XX  
 XX Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;  
 SO

Query Match 91.6%; Score 654.8; DB 18; Length 15222;  
 Best Local Similarity 94.8%; Pred. No. 2e-147;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACATAAAGTCAATGATGACAAACCAAGCCAGATCAAGACACA 60  
 DB 4886 CACAAAGTCACACCAACTGATATCAAGATGCAACAGCCAGATCAAGACACA 4945  
 QY 61 ACCCAACATACCTACTGAGATCTCAGCTTGGAATGAGCTTCTCAATCTGTGAA 120  
 DB 4946 ACCCAACATACCTGACCCAGAAATCTGAGTTGGATTCAGTCCCTCTTAATCCCTGAA 5005  
 QY 121 ATTGATGACAAACCCACCACCTACTGATGATCAACAAACAGGATCAAGTCAACCTG 180  
 DB 5006 ATTGATGACAAACCCACCACCTACTGATGATCAACAAACAGGATCAAGTCAACCTG 5065  
 QY 181 CAACCCACACAGTCAAGAGCTAAAGCAACAAACCAACCAACCAACCAACCAACCCG 240  
 DB 5066 CAATCCACACAGTCAAGAGCAACAAACCAACCAACCAACCAACCAACCAACCCG 5125  
 QY 241 ACTCAAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 DB 5126 ACCCAAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185  
 QY 301 GTGTTAACTTGTACCTGTGAGATATGACAGCAACATCCACCTGCGCTATCTGC 360  
 DB 5186 GTGTTAACTTGTACCTGTGAGATATGACAGCAACATCCACCTGCGCTATCTGC 5245  
 QY 361 AAAAGATATACCAACCAAAACCAAGCAACCAACCAACCAACCAACCAACCAACCA 420  
 DB 5246 AAAAGATATACCAACCAAAACCAAGCAACCAACCAACCAACCAACCAACCAACCA 5305

QY 421 ACCTTCAAGACCAACCAAAAGATCTCAAAACCTCAACCACTAAACCAAGCAATACCC 480  
 DB 5306 ACCCTCAAGACCAACCAAAAGATCTCAAAACCTCAACCACTAAACCAAGCAATACCC 5365  
 QY 481 ACCCAAGCCCAAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB 5366 ACCCAAGCCCAAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCA 5425  
 QY 541 CTGCTACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 DB 5426 CTACTACACCTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5485  
 QY 601 TCACACCTCTCGAGAGCAATCTAAGCCTTCTCAAGTCTCCACCAACATCCGACACCA 660  
 DB 5486 TCACACCTCTCGAGAGCAATCTAAGCCTTCTCAAGTCTCTACACATCCGACACCA 5545  
 QY 661 TCACACCTCTCTACCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 714  
 DB 5546 TCACACCTCTCTACCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 5599

RESULT 9  
 AAT63430  
 ID AAT63430 standard; DNA; 15223 BP.  
 XX  
 XX AAT63430:  
 AC  
 XX 02-JUL-1997 (first entry)  
 DT  
 XX Respiratory syncytial virus anti-genome.  
 DE  
 XX RSV; vaccine; gene therapy; upper respiratory tract disease; ss.  
 KM  
 XX Human respiratory syncytial virus strain A2.  
 OS  
 XX W09712032-A1.  
 PN  
 XX 03-APR-1997.  
 PD  
 XX 27-SEP-1996; 96WO-US15524.  
 PF  
 XX 27-SEP-1995; 95US-0007083.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Collins PL;  
 PI  
 XX WPI: 1997-212893/19.  
 DR  
 XX Infectious respiratory syncytial virus particles - useful for  
 XX treatment of RSV or gene therapy of upper respiratory tract diseases  
 PT  
 XX Claim 46; Page 43-51; 66pp; English.  
 PS  
 XX A human respiratory syncytial virus (RSV) anti-genome sequence  
 CC (AAT63430) is the 5' to 3' positive-sense sequence of RSV; the genome  
 CC itself is negative-sense. It was synthesised in segments by RT-PCR  
 CC using intracellular RSV mRNA or genomic RNA isolated from purified  
 CC viruses as template. Restriction site markers were introduced by  
 CC incorporating the changes into the primers used for RT-PCR. The  
 CC recombinant sequence can be expressed with a nucleocapsid protein,  
 CC a nucleocapsid phosphoprotein, a large polymerase protein and an  
 CC RNA elongation factor to produce isolated infectious RSV particles  
 CC useful for generating vaccines against RSV. Recombinant RSV genome  
 CC or antigenome can also be used as a vector for gene therapy of the  
 CC upper respiratory tract.  
 CC  
 SO Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

Query Match 91.6%; Score 654.8; DB 18; Length 15223;  
 Best Local Similarity 94.8%; Pred. No. 2e-147;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

OY 1 CACAAAGTCACACTACATGCAATCATATGATGCAACAAGCCAGATCAAGAACA 60
   |||||||
Db 4887 CACAAAGTCACACCAACCACTGCATCATATGATGCAACAAGCCAGATCAAGAACA 4946
OY 61 ACCCAACATACCTACCTAGATGCTGAGTTCAGTTCCTCAATCTGTGTGAA 120
   |||||||
Db 4947 ACCCAACATACCTACCTAGATGCTGAGTTCAGTTCCTCAATCTGTGTGAA 5006
OY 121 ATTACATCAACCAACCACTGATGCTGATGCAACAAGCCAGATCAAGTCAACCTG 180
   |||||||
Db 5007 ATTACATCAACCAACCACTGATGCTGATGCAACAAGCCAGATCAAGTCAACCTG 5066
OY 181 CAACCCACAACAGTCAAGCTCAACCAACCAACCAACCAACCAACCAACCAACCC 240
   |||||||
Db 5067 CAATCCACAACAGTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCC 5126
OY 241 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCTT 300
   |||||||
Db 5127 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCTT 5186
OY 301 GTGTTAACTTTGTACCTGCAAGATGCAATGCAACCAACCTGCTGGCTATCTGC 360
   |||||||
Db 5187 GTGTTAACTTTGTACCTGCAAGATGCAATGCAACCAACCTGCTGGCTATCTGC 5246
OY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420
   |||||||
Db 5247 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5306
OY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 480
   |||||||
Db 5307 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 5366
OY 481 ACCACCAAGCCACAGAGAGCCACCAACCAACCAACCAACCAACCAACCAACCT 540
   |||||||
Db 5367 ACCACCAAGCCACAGAGAGCCACCAACCAACCAACCAACCAACCAACCAACCT 5426
OY 541 CTGCTACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTTCC 600
   |||||||
Db 5427 CTGCTACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTTCC 5486
OY 601 TCACACTCTCCGAGAGCAATCTAAGCCTTCTCAAGTCTCCACCAACCTCCAGC 660
   |||||||
Db 5487 TCACACTCTCCGAGAGCAATCTAAGCCTTCTCAAGTCTCCAGCAGTACCA 5546
OY 661 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAAC 714
   |||||||
Db 5547 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAAC 5600

```

```

RESULT 10
AAV17553
ID AAV17553 standard; cDNA; 15223 BP.
AC AAV17553;
XX
XX
XX 20-JUL-1998 (first entry)
XX
XX Respiratory syncytial virus antigenome.
DE
XX
XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.
XX
XX Human respiratory syncytial virus D46.
XX
XX WO9802530-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97WO-US12269.
XX
XX 23-MAY-1997; 97US-0047634.
XX
XX 15-JUL-1996; 96US-0021773.
XX
XX 09-MAY-1997; 97US-0046141.

```

```

XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Bukreyev AA, Collins PL, Juhász K, Murphy BR, Teng MN;
PI Whitehead SS;
XX
XX WPI; 1998-110579/10.
XX
XX Attenuated respiratory syncytial virus vaccines - useful to protect
XX individuals against RSV infection
XX
XX Example 7; Page 188-195; 238pp; English.
XX
XX This is the 5'-3' positive sequence nucleotide sequence of
XX respiratory syncytial virus (RSV) D46. The genome is
XX negative-sense; the complete nucleotide sequence of the
XX wild-type B-1 virus has also been determined (see AAV17552).
XX A novel infectious recombinant RSV comprises a RSV genome or
XX antigenome, a major nucleocapsid (N) protein, a nucleocapsid
XX phosphoprotein (P), a large polymerase protein (L), and a RNA
XX polymerase elongation factor, where the recombinant RSV has at
XX least two attenuating mutations, one of the mutations specifying a
XX temperature-sensitive (ts) substitution at amino acid phe521.
XX Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts
XX nucleotide substitution in the gene-start sequence of gene M2.
XX Also claimed are: (1) an isolated infectious RSV particle which
XX comprises a recombinant RSV (anti)genome, N, P, and L proteins, a
XX RNA polymerase elongation factor, where the (anti)genome is modified:
XX (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or
XX a cis-acting regulatory sequence; and (ii) by a termination codon
XX introduced within a selected gene, or by a change in sequence,
XX position or presence of a GS or GE transcription signal relative to
XX the selected gene; (2) an expression vector; and (3) an RSV strain
XX selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR
XX 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),
XX cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or
XX B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated
XX attenuated recombinant RSV and RSV particles are used in a vaccine
XX to stimulate the immune system of an individual to induce
XX protection against RSV. The expression vector of (2) is used for
XX the production of infectious attenuated RSV particles.
XX
XX Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other:
XX
XX Query Match 91.6%; Score 654.8; DB 19; Length 15223;
XX Best Local Similarity 94.8%; Pred. No. 2e-147;
XX Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

Db 5247 AAAAGATACCAAAAAAACAGGAAAGAAACCACTACCAAGCCCAAAAAAACCA 5306
OY 421 ACCTTAAGACACCAAAAAAGATCTCAAAACCTCAAAACCAACCAAGAACTACCC 480
Db 5307 ACCCTCAAGACCAAAAAAGATCTCAAAACCTCAAAACCAAGAACTACCC 5366
OY 481 ACCACCAAGCCCAAGAGAGCCCAACCAATCAACACCAACCAACCAACCACTACA 540
Db 5367 ACCACCAAGCCCAAGAGAGCCCAACCAATCAACACCAACCAACCAACCACTACA 5426
OY 541 CTGCTCACCACACACACAGAGAAATCCAAATCTCAAGTCAAAATGGAAACCTTCAC 600
Db 5427 CTACTCTACTCCACACACAGAGAAATCCAGAACTCAAAATGGAAACCTTCAC 5486
OY 601 TCACACTCTCCGAAAGGCAATCTAAAGCCCTTCCAAGTCCCAACATCCGAGCACCA 660
Db 5487 TCACACTCTCCGAAAGGCAATCTAAAGCCCTTCCAAGTCCCAACATCCGAGCACCA 5546
OY 661 TCACAAACCTCTCTCCACCAACACACACCCGAGTACTTATTAATAAAAAA 714
Db 5547 TCACAAACCTCTCTCCACCAACACACACCCGAGTACTTATTAATAAAAAA 5600

```

## RESULT 11

AAA88743 standard; cDNA: 15223 BP.

AAA88743;

19-FEB-2001 (first entry)

Respiratory syncytial virus DA6 5'-3' positive sense sequence.

RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.

OS Chimeric - Human respiratory syncytial virus.

OS Chimeric - Bacteriophage T7.

Key Location/Qualifiers

FT variation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

FT mutation

FT 1099

```

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Collins PL, Murphy BR, Whitehead SS;
XX WPI: 2000-679462/66.
XX
XX Infectious chimeric respiratory syncytial virus (RSV) produced from
XX cloned nucleotide sequences, useful as a vaccine against diseases
XX caused by the virus, such as pneumonia and bronchiolitis -
XX
XX Example 7; Page 262-268; 280pp; English.
XX
XX The present sequence is that of the 5' to 3' positive-sense
XX sequence of human respiratory syncytial virus (RSV); the genome
XX itself is negative-sense. This antigenome cDNA, termed DA6, was
XX synthesized in segments by RT-PCR using synthetic oligonucleotides
XX as primers and intracellular RSV mRNA or genome RNA isolated from
XX purified virions as template. The antigenome includes a 5'-terminal
XX nonviral G triplet contributed by the T7 promoter, 4 sequence
XX markers (see AAA8745-47) at positions 1099 (which adds 1 nucleotide
XX to the length), 1139, 5611 and 7559, a ribozyme and tandem T7
XX terminators, and a single nonviral 3'-phosphorylated residue
XX contributed to the 3' end by ribozyme cleavage. The invention
XX provides an isolated infectious chimeric RSV comprising a major
XX nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a
XX large polymerase protein (L), an RNA polymerase elongation factor,
XX and a partial or complete RSV genome or antigenome of one RSV
XX strain or subgroup virus combined with a heterologous gene of a
XX different RSV strain or subgroup virus. The chimeric RSV is
XX infectious and attenuated, preferably by introduction of selected
XX mutations. It is useful as a vaccine against RSV, which causes
XX diseases such as pneumonia and bronchiolitis in infants. The
XX immune system of an individual is stimulated to induce protection
XX against natural RSV infection, preferably in a multivalent manner
XX to achieve protection against multiple RSV strains and/or subgroups.
XX
XX Sequence 15223 BP; 5921 A; 2705 C; 2361 G; 4236 T; 0 other;
XX
XX Query Match 91.6%; Score 654.8; DB 21; Length 15223;
XX Best Local Similarity 94.8%; Pred. No. 2e-147;
XX Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

OY 1 CACAAAGTCACACTAAGCAATCATCAAGATCAACCAAGCCAGATCAAGAACACA 60
Db 4887 CACAAAGTCACACCAACCAAGTCATCAAGATCAACCAAGCCAGATCAAGAACACA 4946
OY 61 ACCCCACATACCTCAGATCAGATCCTCAGCTTGAGTACAGCTTCTGTCTGAA 120
Db 4947 ACCCCACATACCTCAGATCAGATCCTCAGCTTGAGTACAGCTTCTGTCTGAA 5006
OY 121 ATTACATCACAAACCAACCACTATCTCAACCAACCAAGAGTCAAGTCAAACTG 180
Db 5007 ATTACATCACAAACCAACCACTATCTCAACCAACCAAGAGTCAAGTCAAACTG 5066
OY 181 CAACCCACAAAGTCAGACTATAAACAACAACAACCAACCAACCAACCAAGCC 240
Db 5067 CAATCCACAAAGTCAGACTATAAACAACAACAACCAACCAACCAACCAAGCC 5126
OY 241 ACTCAAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5127 ACTCAAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186
OY 301 GTGTTAACTTTGACCTCAGATATGAGAGCAACCAACCAACCAACCAACCAACCA 360
Db 5187 GTGTTAACTTTGACCTCAGATATGAGAGCAACCAACCAACCAACCAACCAACCA 5246
OY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 5247 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306
OY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
Db 5307 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366

```

QY 481 ACCACGAGCCGACAGAGAGCCACCATCAACACGACCAACCAACATCACTACA 540  
|||||  
Db 5367 ACCACGAGCCGACAGAGAGCCACCATCAACACGACCAACCAACATCACTACA 5426  
QY 541 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 600  
|||||  
Db 5427 CTACTCAGCTCCAGACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 5486  
QY 601 TCACACTCTCCGAGAGCAATCAAGCCCTTCCAAAGTCTCCACAAATCCAGACCCA 660  
|||||  
Db 5487 TCACACTCTCCGAGAGCAATCAAGCCCTTCCAAAGTCTCCACAAATCCAGACCCA 5546  
QY 661 TCACAAACCCGATCTCCACGACACACAGCCAGTAGTTATTTAAAAAAA 714  
|||||  
Db 5547 TCACAACTTCTATCTCCACCAACACACAGCCAGTAGTTACTTAAAAACATA 5600

RESULT 12  
AAC88494  
ID AAC88494 standard; DNA; 918 BP.  
XX  
AC AAC88494;  
XX  
DT 12-MAR-2001 (first entry)  
XX  
DE Human RSV G-protein gene.  
XX  
KW Chimeric: respiratory syncytial virus; RSV; immunize; ds.  
XX  
OS Unidentified.  
XX  
PN WO200068392-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 10-MAY-2000; 2000WO-US12582.  
XX  
PR 11-MAY-1999; 99US-0133536.  
XX  
PA (UNIT ) UNIV ILLINOIS FOUND.  
XX  
PI Buelow DE, Korban SS, Sandhu J, Krasnyanski SF;  
XX  
DR WPI; 2001-122707/13.  
XX  
PT Chimeric nucleic acid construct for immunizing animals and humans  
PT against respiratory syncytial virus (RSV), comprises a sequence adapted  
PT for expression in plants and a RSV protein or peptide coding sequence  
PT  
XX  
PS Disclosure; Fig 10; 67pp; English.  
XX  
CC The present invention relates to a chimeric nucleic acid construct  
CC comprising: a nucleotide sequence adapted for protein expression in  
CC plants; and a respiratory syncytial virus (RSV) coding sequence  
CC encoding an RSV protein or an antigenic protein or peptide of RSV.  
CC The construct can be used to immunize animals and humans against  
CC respiratory syncytial virus. The use of transgenic plants to  
CC generate the antigen allows the production of greater amounts of  
CC antigen.  
XX  
SQ Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;

Query Match 90.9%; Score 649.6; DB 22; Length 918;  
Best Local Similarity 95.2%; Pred. No. 1.7e-146;  
Matches 670; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAATGCAATCATATCAAGATGCAACAGCCAGATCAAGACACA 60  
|||||  
Db 214 CACAAAGTCACACTAATGCAATCATATCAAGATGCAACAGCCAGATCAAGACACA 273  
QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGCTTCCCAATCTGTCTGAA 120

Db 274 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGCTTCCCAATCTGTCTGAA 333  
QY 121 ATTACATGCAACACACACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 180  
Db 334 ATTACATGCAACACACACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 333  
QY 181 CAACCCACACAGTCAAGACTTAAAGACACACACCAACCAACCAACCAACCAACCC 240  
Db 394 CAATCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 453  
QY 241 ACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300  
Db 454 ACCCAACAAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 513  
QY 301 GTGTTAACTTTTAACTTGAACCTGACATATGACGAAACCAACCTGCTGGCTATCTGC 360  
Db 514 GTGTTAACTTTTGAACCTGACATATGACGAAACCAACCTGCTGGCTATCTGC 573  
QY 361 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420  
Db 574 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 633  
QY 421 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Db 634 ACCCTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693  
QY 481 ACCCAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 540  
Db 694 ACCCAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 753  
QY 541 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 600  
Db 754 CTACTCAGCTCCAGACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 813  
QY 601 TCACACTCTCCGAGAGCAATCAAGCCCTTCCAAAGTCTCCACAAATCCAGACCCA 660  
Db 814 TCACACTCTCCGAGAGCAATCAAGCCCTTCCAAAGTCTCCACAAATCCAGACCCA 873  
QY 661 TCACAAACCCGATCTCCACGACACACAGCCAGTAGTTATTTAAAAAAA 704  
Db 874 TCACAACTTCTATCTCCACCAACACACAGCCAGTAGTTACT 917

RESULT 13  
AAS20145  
ID AAS20145 standard; DNA; 897 BP.  
XX  
AC AAS20145;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Respiratory syncytial virus G protein DNA.  
XX  
KW RSV; ds; G protein; heavily glycosylated protein; antihaemic; antiviral;  
KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
KW antiviral chemotherapeutic compound; humoral response;  
KW cellular immune response; hPIV; paediatric respiratory disease;  
KW globin gene transfer; sickle cell disease; beta-thalassemia;  
KW human immunodeficiency virus infection; HIV.  
XX  
OS Human respiratory syncytial virus.  
XX  
FH Key Location/Qualifiers  
FH CDS 1..897  
FH FT /\*tag= a  
FT FT /product= "G protein"  
XX  
PN WO200192548-A2.  
XX  
PD 06-DEC-2001.  
XX  
PF 22-MAY-2001; 2001WO-US16610.

```

XX 01-JUN-2000; 2000US-208701P.
PR
XX
PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Portner A, Takimoto T;
XX
DR WPI: 2002-130534/17.
DR P-PSDB: AAU74676.
XX
PT Recombinant Sendai virus useful in vaccines to protect infection by
PT paramyxoviruses, comprises exogenous nucleic acid encoding
PT paramyxovirus protein or its antigenic fragment
XX
PS
XX Disclosure: Page 47; 57pp; English.
XX
XX The invention relates to a recombinant Sendai virus comprising an
XX exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its
XX antigenic fragment. The virus may be administered in combination
XX with an antiviral chemotherapeutic compound. Two or more viruses
XX expressing different PMV proteins may be co-administered. Compositions
XX comprising the virus are useful for eliciting a humoral and/or
XX cellular immune response to a PMV in a mammal, particularly a human.
XX Further a recombinant Sendai virus comprising an exogenous nucleic acid
XX encoding a second PMV protein is also administered and priming and/or
XX boosting humoral or cellular immune response comprises administering
XX one or more of a recombinant or isolated PMV protein or its antigenic
XX fragment, a DNA vaccine encoding the same, and a non-Sendai viral
XX vector encoding a PMV protein. The recombinant virus is useful as an
XX effective vaccine against hpiV or RSV (the major causes of paediatric
XX respiratory disease) and also to express any gene of
XX interest in target cells, providing a positive medical impact on
XX impaired cells. Wild-type globin gene transfer (i.e. gene therapy)
XX into stem cells effects a cure for sickle cell disease or beta-
XX thalassemia. The recombinant virus may also prove effective in
XX conferring immunity to human immunodeficiency virus (HIV) infection.
XX The Sendai virus replicates at level that is high enough to
XX induce sufficient immunity, but does not cause any harm to human
XX recipient. The present sequence encodes a respiratory syncytial
XX virus (RSV) G protein (heavily glycosylated protein), a PMV protein
XX suitable for expression by the recombinant virus of the invention.
XX
XX Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other:
XX
Query Match 90.4%; Score 646.2; DB 24; Length 897;
Best Local Similarity 95.3%; Pred. No. 1.1e-145;
Matches 666; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAGCACTGCAATCAATCAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB 199 CACAAAGTCACACCAACCACTGCAATCAATCAAGATGCAACAAAGCCAGATCAAGAACACA 258
QY 61 ACCCCACATACCTCAGCTAGGATCCTCAGTTGGAATAGCTTCCATATGCTGTGAA 120
DB 259 ACCCCACATACCTCAGCTAGGATCCTCAGTTGGAATAGCTTCCATATGCTGTGAA 318
QY 121 ATTATATCAACCAACCACTAGCTTCAACCAACCAAGGAGTCAAGTCAAAACCTG 180
DB 319 ATTATATCAACCAACCACTAGCTTCAACCAACCAAGGAGTCAAGTCAAAACCTG 378
QY 181 CAACCCACCAAGTCAAGGCTAAAAACCAACCAACCAACCAACCAACCAAGCC 240
DB 379 CAATCCACCAAGTCAAGGCTAAAAACCAACCAACCAACCAACCAACCAAGCC 438
QY 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGCC 300
DB 439 ACCCAACCAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAAGCC 498
QY 301 GTGTTTAACTTTGACCTGCGACATATGACGACCAACCAACCAACCTGCTGGCTATTCG 360
DB 499 GTGTTTAACTTTGACCTGCGACATATGACGACCAACCAACCAACCTGCTGGCTATTCG 558
QY 361 AAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420

```

```

DB 559 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618
QY 421 ACCCTCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 619 ACCCTCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678
QY 481 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 679 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738
QY 541 CTGCTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 739 CTACTTACCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798
QY 601 TCACCTCTCCGAAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACCTCCAGACCA 660
DB 799 TCACCTCTCCGAAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACCTCCAGACCA 858
QY 661 TCACCAACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 699
DB 859 TCACCAACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 897

RESULT 14
AAV38298
ID AAV38298 standard; DNA; 696 BP.
XX
XX AAV38298;
XX
XX 26-OCT-1998 (first entry)
XX
XX Respiratory syncytial virus glycoprotein G gene portion.
XX
XX RSV, glycoprotein G, pneumococcal surface protein A; PspA;
XX infection; Streptococcus pneumoniae; sepsis; otitis media;
XX meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;
XX ss.
XX
XX Respiratory syncytial virus.
XX
XX W09824927-A1.
XX
XX 11-JUN-1998.
XX
XX 04-DEC-1997; 97WO-US22847.
XX
XX 04-DEC-1996; 96US-0759505.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Curiel DT, McDaniel LS;
XX
XX WPI: 1998-333343/29.
XX
XX Example 1; Fig 1C; 47pp; English.
XX
XX This is a portion of the respiratory syncytial virus glycoprotein G
XX (RSVg) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)
XX to create plasmid pGTA. This plasmid contains a human
XX cytomegalovirus immediate early promoter and the RSVg gene portion
XX such that when an in-frame fusion is made, the resultant fusion
XX protein may be transported to, and anchored in, a mammalian cell
XX membrane where it can be exposed to the host immune system.
XX Insertion of pneumococcal surface protein A (PspA) coding sequence
XX created plasmid pKSD2601. Intramuscular immunisation of BALB/c
XX mice with pKSD2601 induced protection against an otherwise lethal
XX challenge with a capsular type 3 pneumococcus. A claimed plasmid
XX for expression of pneumococcal epitope DNA in eukaryotic cells
XX

```

includes a promoter for driving expression in a eukaryotic cell (e.g. HCW-1E), DNA encoding a leader sequence (e.g. of RSV) and DNA encoding a pneumococcal epitope such as PspA. The invention also provides a vaccine comprising the plasmid and a suitable carrier or diluent, and optionally one or more cytokines or DNA encoding them, or a bacterial delivery system. The vaccine is used to elicit an immunological response in a host, including humans, susceptible to pneumococcal infection or sepsis. The plasmid can also be used to express a pneumococcal epitope of interest in vitro.

Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;

Query Match 62.3%; Score 445.8; DB 19; Length 696;  
Best Local Similarity 95.3%; Pred. No. 1.3e-97;  
Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

OY 1 CACAAAGTCACACAACTGATCATATACAGATGCAACAGCCAGATCAAGACACA 60  
DB 199 CACAAAGTCACACAACTGATCATATACAGATGCAACAGCCAGATCAAGACACA 258  
OY 61 ACCCCACATACCTGAGGATCTGAGTGGATGATGCTTCTCATCTGTGAA 120  
DB 259 ACCCCACATACCTGAGGATCTGAGTGGATGATGCTTCTCATCTGTGAA 318  
OY 121 ATTACATCAAAACGACACCATCTACTGCTTCAACACAGAGAGTCAAACTGTG 180  
DB 319 ATTACATCAAAATCACACACCTACTGCTTCAACACAGAGAGTCAAACTGTG 378  
OY 181 CAACCCACAAAGTCAGTCAAGTCAAAACACAAAC- AACCCAAACACAAACGACAGCC 239  
DB 379 CAATCCACAAAGTCAGTCAAGTCAAAACACAAAC- AACCCAAACACAAACGACAGCC 438  
OY 240 CACTACAAACAAAGCCCAAAACAAACCAACCAACCAATATGATTTTCACTTGA 299  
DB 439 CACCAACAAACAAAGCCCAAAACAAACCAACCAATATGATTTTCACTTGA 498  
OY 300 AGTGTTAACCTTGTACCTGAGCATATGACAAACATCCACTGCTGGGCTATCTG 359  
DB 499 AGTGTTAACCTTGTACCTGAGCATATGACAAACATCCACTGCTGGGCTATCTG 558  
OY 360 CAAAGATATCAAAACAAACAAACCAAGAAAGAAACCAACCAAGCTTCAAAACAAAC 419  
DB 559 CAAAGATATCAAAACAAACCAAGAAAGAAACCAACCAAGCTTCAAAACAAAC 618  
OY 420 AACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAAACCAAGAGTACC 479  
DB 619 AACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAAACCAAGAGTACC 678  
OY 480 CACCCACCAAGCC 492  
DB 679 CACCCACCAAGCC 691

## RESULT 15

AAV17552 standard; cDNA: 15225 BP.

AAV17552;

20-JUL-1998 (first entry)

Respiratory syncytial virus genome.

RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.

Human respiratory syncytial virus B-1.

WO9802530-A1.

22-JAN-1998.

15-JUL-1997; 97WO-US12269.

XX

PR 23-MAY-1997; 97US-0047634.  
PR 15-JUL-1996; 96US-0021773.  
PR 09-MAY-1997; 97US-0046141.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Bukreyev AA, Collins PL, Juhász K, Murphy BR, Teng MN;

Whitehead SS;

WPI; 1998-110579/10.

Attenuated respiratory syncytial virus vaccines - useful to protect individuals against RSV infection

Example 4; Page 195-202; 238pp; English.

This is the complete nucleotide sequence of the wild-type B-1 respiratory syncytial virus (RSV). The genome is negative-sense; the 5'-3' positive-sense sequence of D46 is provided in AAV17553. A novel infectious recombinant RSV comprises a RSV genome or antigenome, a major nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a large polymerase protein (L), and a RNA polymerase elongation factor, where the recombinant RSV has at least two attenuating mutations, one of the mutations specifying a temperature-sensitive (ts) substitution at amino acid Phe221, Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts nucleotide substitution in the gene-start sequence of gene M2. Also claimed are: (1) an isolated infectious RSV particle which comprises a recombinant RSV (antigenome, N, P, and L proteins), a RNA polymerase elongation factor, where the (antigenome is modified: (1) to ablate or modulate expression of a SH, NS1, NS2 or G gene or a cis-acting regulatory sequence; and (11) by a termination codon introduced within a selected gene, or by a change in sequence, position or presence of a GS or GE transcription signal relative to the selected gene; (2) an expression vector; and (3) an RSV strain selected from cpts RSV 248 (ATCC VR 2450), cpts RSV 530 (ATCC VR 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2455), or cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated attenuated recombinant RSV and RSV particles are used in a vaccine to stimulate the immune system of an individual to induce protection against RSV. The expression vector of (2) is used for the production of infectious attenuated RSV particles.

Sequence 15225 BP; 5904 A; 2710 C; 2399 G; 4212 T; 0 other;

Query Match 42.0%; Score 300.4; DB 19; Length 15225;  
Best Local Similarity 64.6%; Pred. No. 2.3e-62;  
Matches 464; Conservative 0; Mismatches 251; Indels 3; Gaps 1;

OY 1 CACAAAGTCACACTGACATGCAATATACAGATGCAACAGCCAGATCAAGACACA 60  
DB 488 CACAAAGTCACACTGACATGCAATATACAGATGCAACAGCCAGATCAAGACACA 4947  
OY 61 ACCCCACATACCTGAGGATCTGAGTGGATGATGCTTCTCATCTGTGAA 120  
DB 4948 ACCCCACATACCTGAGGATCTGAGTGGATGATGCTTCTCATCTGTGAA 5007  
OY 121 ATTACATCAAAACGACACCATCTACTGCTTCAACACAGAGAGTCAAACTGTG 180  
DB 5008 ATTACATCAAAACGACACCATCTACTGCTTCAACACAGAGAGTCAAACTGTG 5067  
OY 181 CACCCACCAAGTCAGTCAAGTCAAAACACAAACCAACCAACCAAGAGTACC 240  
DB 5068 CACCCACCAAGTCAGTCAAGTCAAAACACAAACCAACCAACCAAGAGTACC 5127  
OY 241 ACTACAAACAAAGCCCAAAACCAACCAACCAACCAACCAATATGATTTTCACTTGA 300  
DB 5128 ACTACAAACAAAGCCCTTAATAATCCACCAAAACCAACCAAGATGATTTTGA 5187  
OY 301 GTGTTTAACCTTGTACCTGAGCATATGACAAACATCCACTGCTGGGCTATCTGC 360  
DB 5188 GTGTTTAACCTTGTACCTGAGCATATGACAAACATCCACTGCTGGGCTATCTGC 5247







GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:49:02 ; Search time 30.1743 Seconds  
(without alignments)  
7266.906 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaagtcacactaacac.....gtagtataaaaaaaaaa 715

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	920	2	US-08-467-963C-7
2	715	100.0	920	2	US-08-836-189D-7
3	715	100.0	920	2	US-08-852-344D-7
4	715	100.0	920	3	US-08-344-639E-7
5	715	100.0	920	4	US-08-467-969A-7
6	715	100.0	920	4	US-08-467-961A-7
7	715	100.0	920	4	US-08-001-554A-7
8	696	97.3	894	2	US-08-467-963C-28
9	696	97.3	894	2	US-08-836-189D-28
10	696	97.3	894	3	US-08-852-344D-28
11	696	97.3	894	3	US-08-344-639E-28
12	654.8	91.6	15222	2	US-08-801-898A-23
13	654.8	91.6	15222	4	US-08-962-690-12
14	654.8	91.6	15223	2	US-08-892-403A-1
15	654.8	91.6	15223	4	US-08-720-132-1
16	600.4	42.0	15225	2	US-08-892-403A-2
17	212.8	29.8	1050	3	US-08-836-501-76
18	210.8	29.5	303	3	US-08-721-979A-1
19	210.8	29.5	303	3	US-08-836-504A-1
20	210.8	29.5	303	3	US-08-836-501-1
21	210.8	29.5	303	3	US-09-654-289-1
22	209.6	29.3	1071	4	US-08-836-501-77
23	207.6	29.0	303	3	US-08-721-979A-3
24	207.6	29.0	303	3	US-08-836-504A-3
25	207.6	29.0	303	3	US-08-836-501-3
26	207.6	29.0	303	3	US-09-654-289-3
27	201.2	28.1	303	4	US-08-721-979A-14

28	201.2	28.1	303	3	US-08-836-501-14	Sequence 14, Appl
29	201.2	28.1	303	4	US-09-654-289-14	Sequence 14, Appl
30	128.6	18.0	183	3	US-08-721-979A-23	Sequence 23, Appl
31	128.6	18.0	183	3	US-08-836-501-23	Sequence 23, Appl
32	128.6	18.0	183	4	US-09-654-289-23	Sequence 23, Appl
33	126	17.6	303	3	US-08-721-979A-2	Sequence 2, Appl
34	126	17.6	303	3	US-08-836-504A-2	Sequence 2, Appl
35	126	17.6	303	3	US-08-836-501-2	Sequence 2, Appl
36	126	17.6	303	4	US-09-654-289-2	Sequence 2, Appl
37	125.4	17.5	183	3	US-08-721-979A-30	Sequence 30, Appl
38	125.4	17.5	183	3	US-08-836-501-30	Sequence 30, Appl
39	125.4	17.5	183	4	US-09-654-289-30	Sequence 30, Appl
40	124.2	17.4	177	3	US-08-721-979A-24	Sequence 24, Appl
41	124.2	17.4	177	3	US-08-836-501-24	Sequence 24, Appl
42	124.2	17.4	177	4	US-09-654-289-24	Sequence 24, Appl
43	122.8	17.2	303	3	US-08-721-979A-4	Sequence 4, Appl
44	122.8	17.2	303	3	US-08-836-504A-4	Sequence 4, Appl
45	122.8	17.2	303	3	US-08-836-501-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-467-963C-7  
Sequence 7, Application US/08467963C  
Patent No. 5968776  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A CHIMERIC PROTEIN WHICH CONTERS PROTECTION AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,963C  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/836,189  
FILING DATE: 16-APR-1997  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-467-963C-7

Query Match 100.0%; Score 715; DB 2; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-185;  
 Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 Db CACAAGTCACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 265  
 QY 61 ACCCAACATACCTCACTGAGATCCTGAGCTTGGAAATAGCTTCCATCTGTGAA 120  
 Db ACCCAACATACCTCACTGAGATCCTGAGCTTGGAAATAGCTTCCATCTGTGAA 325  
 QY 121 ATTACATCAACAACACACACTACTAGCTTCAACAACACAGAGATCAAGTCAACCTG 180  
 Db ATTACATCAACAACACACACTACTAGCTTCAACAACACAGAGATCAAGTCAACCTG 385  
 QY 326 ATTACATCAACAACACACACTACTAGCTTCAACAACACAGAGATCAAGTCAACCTG 385  
 Db 181 CAACCCACAAGTCAAGACTTAAACACACAACACCAACCCACAACCCAGCAAGCC 240  
 Db 386 CAACCCACAAGTCAAGACTTAAACACACAACACCAACCCACAACCCAGCAAGCC 445  
 QY 241 ACACCAAAACAGCCAAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 300  
 Db 446 ACACCAAAACAGCCAAACCAACCAACCAACCAACCAATATATATTTTCACTTGAA 505  
 QY 301 GTGTTTAACTTTGTACCTGACAGATATGAGCAACATCAACCTGTGGCTATCTGC 360  
 Db 506 GTGTTTAACTTTGTACCTGACAGATATGAGCAACATCAACCTGTGGCTATCTGC 565  
 QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 Db 566 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGATGCC 480  
 Db 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGATGCC 685  
 QY 481 ACCACCAAGCCACAGAGAGCCACATACACCAACCAACCAACCAACCAACCTAGCA 540  
 Db 686 ACCACCAAGCCACAGAGAGCCACATACACCAACCAACCAACCAACCAACCTAGCA 745  
 QY 541 CTGCTACCAACAACACACAGAGAAATCCAAAACCTCAAGTAATGGAACCTTCCAC 600  
 Db 746 CTGCTACCAACAACACACAGAGAAATCCAAAACCTCAAGTAATGGAACCTTCCAC 805  
 QY 601 TCACACCTCCCGAAGGCAATTAAGCCCTTCAAGTCCGACAATCCGAGCAACCA 660  
 Db 806 TCACACCTCCCGAAGGCAATTAAGCCCTTCAAGTCCGACAATCCGAGCAACCA 865  
 QY 661 TCACACCTCCATCTCCACCAACACACACAGCCAGTAGTTATTTAAAAA 715\*  
 Db 866 TCACACCTCCATCTCCACCAACACACACAGCCAGTAGTTATTTAAAAA 920

## RESULT 2

US-08-838-189D-7  
 ; Sequence 7, Application US/08838189D  
 ; Patent No. 5998169

GENERAL INFORMATION:  
 APPLICANT: KLEIN, Michel H  
 APPLICANT: DU, Run-Pan  
 TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
 TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
 TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
 NUMBER OF SEQUENCES: 38  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: 6th Floor, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/838,189D  
 FILING DATE: 16-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/001,554  
 FILING DATE: 06-JAN-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 9200117.1  
 FILING DATE: 06-JAN-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: STEWART, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-838-189D-7

Query Match 100.0%; Score 715; DB 2; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-185;  
 Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 Db CACAAGTCACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 265  
 QY 61 ACCCAACATACCTCACTGAGATCCTGAGCTTGGAAATAGCTTCCATCTGTGAA 120  
 Db 266 ACCCAACATACCTCACTGAGATCCTGAGCTTGGAAATAGCTTCCATCTGTGAA 325  
 QY 121 ATTACATCAACAACACACACTACTAGCTTCAACAACACAGAGATCAAGTCAACCTG 180  
 Db 326 ATTACATCAACAACACACACTACTAGCTTCAACAACACAGAGATCAAGTCAACCTG 385  
 QY 181 CAACCCACAAGTCAAGACTTAAACACACAACCAACCAACCAACCAACCAAGATGCC 240  
 Db 386 CAACCCACAAGTCAAGACTTAAACACACAACCAACCAACCAACCAACCAAGATGCC 445  
 QY 241 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 300  
 Db 446 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 505  
 QY 301 GTGTTTAACTTTGTACCTGACAGATATGAGCAACATCAACCTGTGGCTATCTGC 360  
 Db 506 GTGTTTAACTTTGTACCTGACAGATATGAGCAACATCAACCTGTGGCTATCTGC 565  
 QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 Db 566 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGATGCC 480  
 Db 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGATGCC 685  
 QY 481 ACCACCAAGCCACAGAGAGCCACATCAACCAACCAACCAACCAACCAACCTAGCA 540  
 Db 686 ACCACCAAGCCACAGAGAGCCACATCAACCAACCAACCAACCAACCAACCTAGCA 745  
 QY 541 CTGCTACCAACAACACACAGAGAAATCCAAAACCTCAAGTAATGGAACCTTCCAC 600  
 Db 746 CTGCTACCAACAACACACAGAGAAATCCAAAACCTCAAGTAATGGAACCTTCCAC 805

QY	Db	QY	Db
601	TCAACCTCTCCGGAAGGCATCTCAAGCCCTCTTCAAGTCTCCACCAACATCCGAGACCCA	661	TCACAAACCCCTCATCTCCACCCCAACACACACACGCGCAGTAGTATTTAAAAA
602		662	
603		663	
604		664	
605		665	
606	TCAACCTCTCTCCGGAAGGCATCTCAAGCCCTCTTCAAGTCTCCACCAACATCCGAGACCCA	666	TCACAAACCCCTCATCTCCACCCCAACACACACACGCGCAGTAGTATTTAAAAA
607		667	
608		668	
609		669	
610		670	
611		671	
612		672	
613		673	
614		674	
615		675	
616		676	
617		677	
618		678	
619		679	
620		680	
621		681	
622		682	
623		683	
624		684	
625		685	
626		686	
627		687	
628		688	
629		689	
630		690	
631		691	
632		692	
633		693	
634		694	
635		695	
636		696	
637		697	
638		698	
639		699	
640		700	
641		701	
642		702	
643		703	
644		704	
645		705	
646		706	
647		707	
648		708	
649		709	
650		710	
651		711	
652		712	
653		713	
654		714	
655		715	
656		716	
657		717	
658		718	
659		719	
660		720	
661		721	
662		722	
663		723	
664		724	
665		725	
666		726	
667		727	
668		728	
669		729	
670		730	
671		731	
672		732	
673		733	
674		734	
675		735	
676		736	
677		737	
678		738	
679		739	
680		740	
681		741	
682		742	
683		743	
684		744	
685		745	
686		746	
687		747	
688		748	
689		749	
690		750	
691		751	
692		752	
693		753	
694		754	
695		755	
696		756	

### RESULT 3

```

US-08-852-344D-7
: Sequence 7, Application US/08852344D
: Patent No. 6017539
:
: GENERAL INFORMATION:
: APPLICANT: KLEIN, Michel H
: APPLICANT: DU, Run-Pan
: APPLICANT: EMASYSHYN, Mary E
: TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
: TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
: TITLE OF INVENTION: SYNCYTIAL VIRUS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6th Floor, 330 University Avenue
: City: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/852,344D
: FILING DATE: 07-MAY-1997
: CLASSIFICATION: 424
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,639
: FILING DATE: 14-NOV-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9200117.1
: FILING DATE: 06-JAN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-668 MIS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
:
: INFORMATION FOR SEO ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 920 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-852-344D-7

```

Query Match	100.0%;	Score 715;	DB 3;	Length 920;
Best Local Similarity	100.0%;	Pred. No. 1.le-185;		
Matches 715;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	CACAAAGTCACGTAACACACTGCAATCATACAAGTGTGCACAAAGCCAAATCAAGACACA	60		
Db 206	CACAAAGTCACGTAACACACTGCAATCATACAAGTGTGCACAAAGCCAAATCAAGACACA	265		
QY 61	ACCCCAACATACCTCCTACTCAGAGATCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTAA	120		
Db 266	ACCCCAACATACCTCCTACTCAGAGATCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTAA	325		
QY 121	ATTTCATCACAACCCAGCAGCATCTAGCTTCAACACACAGAGTCAAGTCAAACTCG	180		

Db	326	ATTAGATGCAAAACACACACATCTAGCTTGAACAAACACACAGAGATGCAAGTCAAACTG	385
QY	181	CAACCCACACACATGTCAGACTTAAAAACACACAACAACCCAAACACACCCAGCAAGCC	240
Db	386	CAACCCACACACAGTCAGAGACTTAAAAACACAACAACCCAAACACACACCCAGAGCC	445
QY	241	ACTACAAAAACAAGCCGCAAAAAACAAACCAACAAACCCAAATATGATTTTCCTTGAA	300
Db	446	ACTACAAAAACAAGCCGCAAAAAACCAACCAACCAACCCAAATATGATTTTCCTTGAA	505
QY	301	GTCGTTTAACCTTGTACCCCTGCAGATATGAGCAACCAATCAACCTGCTGGCTATCTGC	360
Db	506	GTCGTTTAACCTTGTACCCCTGCAGATATGAGCAACCAATCAACCTGCTGGCTATCTGC	565
QY	361	AAAAAGATATCCAAACAAAAAACCGAGAAAGAAAAACCAACCAAGCCCTACAAAPAAAA	420
Db	566	AAAAAGATATCCAAACAAAAAACCGAGAAAGAAAAACCAACCAAGCCCTACAAAPAAAA	625
QY	421	ACCTTCAAGACACCAAAAAAGATCTCAACCTGCAACCACTAAACCAAGAAATACCC	480
Db	626	ACCTTCAAGACACCAAAAAAGATCTCAACCACTCAACCACTAAACCAAGAAATACCC	685
QY	481	ACCAACCAAGCCCAAGAAAGCCCAACCATCAACCAACCAACCAACCAATCAACATACA	540
Db	686	ACCAACCAAGCCCAAGAAAGCCCAACCATCAACCAACCAACCAACCAATCAACATACA	745
QY	541	CTGCTCAACACACACACACACAGAAATCTCAAAATCTCAAGTCAAAATGGAACCTTCCAC	600
Db	746	CTGCTCAACACACACACACACAGAAATCTCAAAATCTCAAGTCAAAATGGAATACCTTCCAC	805
QY	601	TCAACCTTCCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAACATCTGAGCAACCA	660
Db	806	TCAACCTTCCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAACATCTGAGCAACCA	865
QY	661	TCAACAACCTCTATCTCCACACCAACACAACAACGCGAGTAGTATTATTAACAAAAAAA	715
Db	866	TCAACAACCTCTATCTCCACACCAACACAACAACGCGAGTAGTATTATTAACAAAAAAA	920

## RESULT 4

US-08-344-639E-7  
Sequence 7, Application US/08344639E  
Patent No. 6033668  
GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: Ewaszyshyn, Mary E  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/344,639E  
FILING DATE: 14-NOV-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-391 MIS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-344-639E-7

Query Match 100.0%; Score 715; DB 3; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.1e-185;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTACACTGCATCAACAGATGCAACCAAGATCAAGAACACA 60  
DB 206 CACAAAGTCACACTACACTGCATCAACAGATGCAACCAAGATCAAGAACACA 265  
QY 61 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCGATCTGTGAA 120  
DB 266 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCGATCTGTGAA 325  
QY 121 ATTACATCAAAACACACACCTACTAGTCTTCAACCAACACGAGTCAAGTCAACCTG 180  
DB 326 ATTACATCAAAACACACACCTACTAGTCTTCAACCAACACGAGTCAAGTCAACCTG 385  
QY 181 CAACCCCAACAGTCAGATGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 240  
DB 386 CAACCCCAACAGTCAGATGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 445  
QY 241 ACTACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300  
DB 446 ACTACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505  
QY 301 GTGTTTAACTTTGTACCTCAGCATATGACGACCAACCAACCAACCAACCAACCAACCAAC 360  
DB 506 GTGTTTAACTTTGTACCTCAGCATATGACGACCAACCAACCAACCAACCAACCAACCAAC 565  
QY 361 AAAAGATACCAAAACCAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 566 AAAAGATACCAAAACCAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
QY 421 ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAAAGAAAGTACCC 480  
DB 626 ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAAAGAAAGTACCC 685  
QY 481 ACCACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
DB 686 ACCACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 745  
QY 541 CTGCTCAGCAACACCAAGAAATCCAAACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 600  
DB 746 CTGCTCAGCAACACCAAGAAATCCAAACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 805  
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCAACATCCGAGCAACCA 660  
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCAAGTCTCAACATCCGAGCAACCA 865  
QY 661 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715  
DB 866 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

RESULT 5

US-08-467-969A-7  
Sequence 7, Application US/08467969A  
Patent No. 6168786  
GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: Ewasysbyn, Mary E  
TITLE OF INVENTION: Chimeric Immunogens  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,969A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-475 MIS:bn  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-467-969A-7

Query Match 100.0%; Score 715; DB 4; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.1e-185;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTACACTGCATCAACAGATGCAACCAAGATCAAGAACACA 60  
DB 206 CACAAAGTCACACTACACTGCATCAACAGATGCAACCAAGATCAAGAACACA 265  
QY 61 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCGATCTGTGAA 120  
DB 266 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCGATCTGTGAA 325  
QY 121 ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180  
DB 326 ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTG 385  
QY 181 CAACCCCAACAGTCAGATGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
DB 386 CAACCCCAACAGTCAGATGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 445  
QY 241 ACTACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 446 ACTACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505

```

QY 301 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAACATCCACCTGGCTATCTGC 360
    |||||||
Db 506 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAACATCCACCTGGCTATCTGC 565
QY 361 AAAAGATATCGAAGAAAAACAGGAAAAAGAAACACACCAAGCTTACAAAAACCA 420
    |||||||
Db 566 AAAAGATATCGAAGAAAAACAGGAAAAAGAAACACACCAAGCTTACAAAAACCA 625
QY 421 ACCTTCAAGACAAACAAAAAATCTCAAACTTCAAACTTCAAACTTCAAACTTCA 480
    |||||||
Db 626 ACCTTCAAGACAAACAAAAAATCTCAAACTTCAAACTTCAAACTTCAAACTTCA 685
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAACATCAACATCA 540
    |||||||
Db 686 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAACATCAACATCA 745
QY 541 CTGCTCACCACACACACACCAAGAAATCCAAAACTCACAAGTCAAAATGAAAACTTCAC 600
    |||||||
Db 746 CTGCTCACCACACACACCAAGAAATCCAAAACTCACAAGTCAAAATGAAAACTTCAC 805
QY 601 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAACTCTCAACATCCGAGACCA 660
    |||||||
Db 806 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAACTCTCAACATCCGAGACCA 865
QY 661 TCACACCCCTATCTCACCACCAACACACGCGAGTAGTATTAAAAA 715
    |||||||
Db 866 TCACACCCCTATCTCACCACCAACACACGCGAGTAGTATTAAAAA 920

```

RESULT 6  
US-08-467-961A-7

```

; Sequence 7, Application US/08467961A
; Patent No. 6171783
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasashyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 330 University Avenue, 6TH Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467, 961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 7:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-467-961A-7

```

Query Match 100.0%; Score 715; DB 4; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1,1e-185;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CACAAGTGCACATTAACATCGAATCATATCAAGATGCAACGACGATCAAGACACA 60
    |||||||
Db 206 CACAAGTGCACATTAACATCGAATCATATCAAGATGCAACGACGATCAAGACACA 265
QY 61 ACCCAACATACCTCTACTCAGATCTCTCAGCTTGGATTCAGTCTTCGAA 120
    |||||||
Db 266 ACCCAACATACCTCTACTCAGATCTCTCAGCTTGGATTCAGTCTTCGAA 325
QY 121 ATTACATCAACACACACACCATTAAGTTCACAAACACACGAGATCAAACTTG 180
    |||||||
Db 326 ATTACATCAACAAACACACCATTAAGTTCACAAACACACGAGATCAAACTTG 385
QY 181 CAACCCCAACAGTCAAGACATAAAAAACAAACAAACCAACCAACCAACCAAGCC 240
    |||||||
Db 386 CAACCCCAACAGTCAAGACATAAAAAACAAACAAACCAACCAACCAACCAAGCC 445
QY 241 ACTACAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
    |||||||
Db 446 ACTACAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAACATCCACCTGGCTATCTGC 360
    |||||||
Db 506 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAACATCCACCTGGCTATCTGC 565
QY 361 AAAAGATATCGAAGAAAAACAGGAAAAAGAAACACACCAAGCTTACAAAAACCA 420
    |||||||
Db 566 AAAAGATATCGAAGAAAAACAGGAAAAAGAAACACACCAAGCTTACAAAAACCA 625
QY 421 ACCTTCAAGACAAACAAAAAAGATCTCAAACTCAAAACCACTAAACGAAGTACCC 480
    |||||||
Db 626 ACCTTCAAGACAAACAAAAAAGATCTCAAACTCAAAACCACTAAACGAAGTACCC 685
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAACATCAACATCA 540
    |||||||
Db 686 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAACATCAACATCA 745
QY 541 CTGCTCACCACACACACCAAGAAATCCAAAACTCACAAGTCAAAATGAAACCTTCAC 600
    |||||||
Db 746 CTGCTCACCACACACCAAGAAATCCAAAACTCACAAGTCAAAATGAAACCTTCAC 805
QY 601 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAACTTCAAACTTCAAACTTCA 660
    |||||||
Db 806 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAACTTCAAACTTCAAACTTCA 865
QY 661 TCACACCCCTATCTCACCACCAACACACGCGAGTAGTATTAAAAA-AAAA 715
    |||||||
Db 866 TCACACCCCTATCTCACCACCAACACACGCGAGTAGTATTAAAAA 920

```

RESULT 7

```

; Sequence 7, Application US/08001554A
; Patent No. 6225091
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasashyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney

```

STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentlin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/001,554A  
FILING DATE: 06-JAN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-286  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-001-554A-7

Query Match 100.0%; Score 715; DB 4; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1,1e-185;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCAACAAGCCCAATCAAGAACA 60  
DB 206 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCAACAAGCCCAATCAAGAACA 265  
QY 61 ACCCCAACTACTCTCAGAGATCCTGAGTTCGATTCGATTCGATTCGATTCGATTCG 120  
DB 266 ACCCCAACTACTCTCAGAGATCCTGAGTTCGATTCGATTCGATTCGATTCGATTCG 325  
QY 121 ATTACATCAACAACCAACCATCTAGTCTTCAACAACAACCAAGAGTCAAGTCAAACTG 180  
DB 326 ATTACATCAACAACCAACCATCTAGTCTTCAACAACAACCAAGAGTCAAGTCAAACTG 385  
QY 181 CAACCCCAACGCTCAAGACTTAAACACAACAACCAACCAACCAACCAACCAACCAAC 240  
DB 386 CAACCCCAACGCTCAAGACTTAAACACAACAACCAACCAACCAACCAACCAACCAAC 445  
QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 446 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
QY 301 GTGTTTAACCTTTGACCTGACGATATGACGACAAATCCAACTGCTGGGCTATCTGC 360  
DB 506 GTGTTTAACCTTTGACCTGACGATATGACGACAAATCCAACTGCTGGGCTATCTGC 565  
QY 361 AAAAGATACCAACAAAACCAAGGAAAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 566 AAAAGATACCAACAAAACCAAGGAAAACCAACCAACCAACCAACCAACCAACCAACCA 625  
QY 421 ACCTTCAGACAAACAAAAGATCTCAAACTCAAAACCAATTAACCAAGGAGTACCC 480  
DB 626 ACCTTCAGACAAACAAAAGATCTCAAACTCAAAACCAATTAACCAAGGAGTACCC 685  
QY 481 ACCACCAAGCCACAGAAAGCCAAACCATCAACGACCAACCAACCAACCAACCAACCA 540  
DB 686 ACCACCAAGCCACAGAAAGCCAAACCATCAACGACCAACCAACCAACCAACCAACCA 745  
QY 541 CTGCTACCAACAGACAGAGGAATCCAAAACCTCACAAGTCAATTGGAACCTTCCAC 600

DB 746 CTGCTACCAACAGACAGAGGAATCCAAAACCTCACAAGTCAATTGGAACCTTCCAC 805  
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 660  
DB 806 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 865  
QY 661 TCACAAACCTCATCTCCACCAACACACACAGCCAGTAGTTATTAATAAAAAA 715  
DB 866 TCACAAACCTCATCTCCACCAACACACACAGCCAGTAGTTATTAATAAAAAA 920

## RESULT 8

US-08-467-963C-28  
Sequence 28, Application US/08467963C  
Patent No. 5968776

## GENERAL INFORMATION:

APPLICANT: KLEIN, Michel H

APPLICANT: DU, Run-Pan

TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST

TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentlin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,963C

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/838,189

FILING DATE: 16-Apr-1997

APPLICATION NUMBER: US 08/001,554

FILING DATE: 06-JAN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 894 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-467-963C-28

Query Match 97.3%; Score 696; DB 2; Length 894;

Best Local Similarity 100.0%; Pred. No. 1,6e-180; Mismatches 0; Indels 0; Gaps 0;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCAACAAGCCCAATCAAGAACA 60  
DB 199 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCAACAAGCCCAATCAAGAACA 258  
QY 61 ACCCCAACTACTCTCAGAGATCCTGAGTTCGATTCGATTCGATTCGATTCGATTCG 120  
DB 259 ACCCCAACTACTCTCAGAGATCCTGAGTTCGATTCGATTCGATTCGATTCGATTCG 318





Sequence 28, Application US/08852344D  
Patent No. 6017539  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
APPLICANT: EWASYSHYN, Mary E  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,344D  
FILING DATE: 07-MAY-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/344,639  
FILING DATE: 14-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-688 MIS-Jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-852-344D-28  
Query Match 97.3%; Score 696; DB 3; Length 894;  
Best Local Similarity 100.0%; Pred. No. 1.6e-180;  
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 499 GTGTTAACTTTGTACCTGCGATATGAGCAACAATCCAACTGCTGAGCTATCTGCG 558  
QY 361 AAAGAAATACCAACCAAAAAACCGAAGAAACCCACCGAAGCTTACAAAAACCA 420  
Db 559 AAAAGAAATACCAACCAAAAAACCGAAGAAACCCACCGAAGCTTACAAAAACCA 618  
QY 421 ACCTTCAAGCAACCAAAAAAGATCTCAAACTTCAAACTTCAAACTTCAAACTTCA 480  
Db 619 ACCTTCAAGCAACCAAAAAAGATCTCAAACTTCAAACTTCAAACTTCAAACTTCA 678  
QY 481 ACCACCAAGCCGACGAGAGCCGACCATATACACCCCAACCAACCAATCATCTACTACA 540  
Db 679 ACCACCAAGCCGACGAGAGCCGACCATATACACCCCAACCAACCAATCATCTACTACA 738  
QY 541 CTGCTCACCACCAACACACACAGAAATCCAAATCCAAATCCAAATCCAAATCCAAAT 600  
Db 739 CTGCTCACCACCAACACACAGAAATCCAAATCCAAATCCAAATCCAAATCCAAAT 798  
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 660  
Db 799 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 858  
QY 661 TCAACCTCTCTATCTCCACCCCAACCAACGAGCCAG 696  
Db 859 TCACAACCTCTATCTCCACCCCAACCAACGAGCCAG 894

RESULT 11  
US-08-344-639E-28  
Sequence 28, Application US/08344639E  
Patent No. 603568  
GENERAL INFORMATION:  
APPLICANT: Klein, Michel H  
APPLICANT: Du, Run-Pan  
APPLICANT: EWASYSHYN, Mary E  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 330 University Avenue, 6th Floor  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/344,639E  
FILING DATE: 14-NOV-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-391 MIS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
TELEX: 065-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-344-639E-28

```

```

Query Match          97.3%; Score 696; DB 3; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.6e-180;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CACAAAGTCACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 60
DB CACAAAGTCACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 258
QY 61 ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 120
DB ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 318
QY 121 ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 180
DB ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 378
QY 319 ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 378
DB ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 378
QY 181 CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438
QY 379 CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438
DB CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438
QY 241 ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
DB ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 498
QY 439 ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 498
DB ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 498
QY 301 GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 360
DB GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 558
QY 499 GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 558
DB GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 558
QY 361 AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420
DB AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 618
QY 559 AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 618
DB AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 618
QY 421 ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480
DB ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 678
QY 619 ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 678
DB ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 678
QY 481 ACCCAACAAGCCCAAGAAAGCCCAACATCAACACCAACCAACCAACCAACATCAACA 540
DB ACCCAACAAGCCCAAGAAAGCCCAACATCAACACCAACCAACCAACCAACATCAACA 738
QY 679 ACCCAACAAGCCCAAGAAAGCCCAACATCAACACCAACCAACCAACCAACATCAACA 738
DB ACCCAACAAGCCCAAGAAAGCCCAACATCAACACCAACCAACCAACCAACATCAACA 738
QY 541 CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAA 600
DB CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAA 798
QY 739 CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAA 798
DB CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAA 798
QY 601 TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAGCA 660
DB TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAGCA 858
QY 799 TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAGCA 858
DB TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAGCA 858
QY 661 TCACAACCCCTATCTCCACCAACCAACCAACGCGCAG 696
DB TCACAACCCCTATCTCCACCAACCAACCAACGCGCAG 894

```

```

RESULT 12
US-08-801-898A-23
; Sequence 23; Application US/08801898A
; Patent No. 5998602
; GENERAL INFORMATION:
; APPLICANT: Torrence, Paul F.
; APPLICANT: Cirtino, Robert H.
; APPLICANT: Li, Gulying
; APPLICANT: Xiao, Wei
; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,898A
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8656-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15222 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: RSV-A2
; LOCATION: 1..15222
; OTHER INFORMATION:
US-08-801-898A-23

```

```

Query Match          91.6%; Score 654.8; DB 2; Length 15222;
Best Local Similarity 94.8%; Pred. No. 8.5e-169;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

QY 1 CACAAAGTCACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 60
DB CACAAAGTCACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 4945
QY 61 ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 120
DB ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 5005
QY 4946 ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 5005
DB ACCCAACATACCTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 5005
QY 121 ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 180
DB ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 5065
QY 5066 ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 5065
DB ATTACATCAACAACGACCCATCTACTAGCTTCAACAACGACGAGTCAAGTCAAACTG 5065
QY 181 CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB CACCCCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5125
QY 5066 CAAATCCACAACCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5125
DB CAAATCCACAACCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5125
QY 241 ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
DB ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5185
QY 5126 ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5185
DB ACTACAAACAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5185
QY 5186 GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 5245
DB GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 5245
QY 301 GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 360
DB GTGTTTAACCTTTGATCCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 5245
QY 181 AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420
DB AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5305
QY 5246 AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5305
DB AAAAGAAATACCAAAACAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5305
QY 421 ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480
DB ACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 5365
QY 5306 ACCCTTAACAACAACCAAAAGATCTCCAACTCAAACTCAAACTCAAACTCAAACTCA 5365
DB ACCCTTAACAACAACCAAAAGATCTCCAACTCAAACTCAAACTCAAACTCAAACTCA 5365
QY 481 ACCCAACAAGCCCAAGAAAGCCCAACATCAACACCAACCAACCAACATCAACA 540

```





Oy 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCCA 660  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 5487 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGTACCCA 5546  
Oy 661 TCACACCCCTCATCTTCACACCCACACACAGCCAGTAGTTATTAAAAAAA 714  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 5547 TCACACCTTCATCTCCACCCACACACAGCCAGTAGTTACTTAAAAACATA 5600

Search completed: May 1, 2003, 12:35:31  
Job time : 53.1743 secs

GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 11:59:58 ; Search time 73.4679 Seconds  
(Without alignments)  
11499.705 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaagaatcacactaacac.....gtagtattataaaaaaaa 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCF\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCFUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	654.8	91.6	15223	9 US-09-847-173-1	Sequence 1, Appli
2	445.8	62.3	696	10 US-09-844-645-4	Sequence 4, Appli
3	83.6	11.7	793	10 US-09-878-574-4304	Sequence 4304, Ap
4	76.6	10.7	576	10 US-09-878-574-4296	Sequence 4296, Ap
5	76.4	10.7	529	10 US-09-883-965-2109	Sequence 2109, Ap
6	75.6	10.6	1030	10 US-09-878-574-4306	Sequence 4306, Ap
7	75	10.5	577	10 US-09-878-574-4312	Sequence 4312, Ap
8	74	10.3	639	10 US-09-878-574-4316	Sequence 4316, Ap
9	71.4	10.0	1075	10 US-09-864-761-19241	Sequence 19241, A
10	71.4	10.0	1403	10 US-09-864-761-2513	Sequence 2513, Ap
11	71.2	10.0	1635	10 US-09-864-761-20241	Sequence 20241, A
12	71.2	10.0	1973	10 US-09-864-761-3471	Sequence 3471, Ap
13	70	9.8	576	10 US-09-864-761-26582	Sequence 26582, A
14	68.6	9.6	1423	9 US-10-001-857-100	Sequence 100, App
15	68.2	9.5	491	10 US-09-878-574-4301	Sequence 4301, Ap
16	67.6	9.5	584	10 US-09-878-574-4310	Sequence 4310, Ap
17	67.2	9.4	1428	9 US-10-001-857-79	Sequence 79, Appl
18	63.6	8.9	555	10 US-09-864-761-9605	Sequence 9605, Ap
19	63.4	8.9	766	10 US-09-878-574-4344	Sequence 4344, Ap

C 20	63	8.8	528	10	US-09-878-574-4305	Sequence 4305, Ap
C 21	63	8.8	1267	12	US-10-001-843-45	Sequence 45, Appl
C 22	61.2	8.6	446	10	US-09-960-352-3400	Sequence 3400, Ap
C 23	61	8.5	448	10	US-09-878-574-4308	Sequence 4308, Ap
C 24	60.2	8.4	341	10	US-09-960-352-12302	Sequence 12302, A
C 25	59	8.3	545	10	US-09-878-574-4299	Sequence 4299, Ap
C 26	57.6	8.1	1557	9	US-10-001-873-1	Sequence 1, Appl1
C 27	57.4	8.0	425	10	US-09-960-352-4010	Sequence 4010, Ap
C 28	57.4	8.0	1938	9	US-09-834-975-1039	Sequence 1039, Ap
C 29	57.2	8.0	11956	9	US-10-270-336-3	Sequence 3, Appl1
C 30	57	8.0	446	10	US-09-864-761-20699	Sequence 20699, A
C 31	56	7.8	155074	9	US-10-026-188-6	Sequence 6, Appl1
C 32	55.6	7.8	283	10	US-09-960-352-9095	Sequence 9095, Ap
C 33	54	7.6	415	10	US-09-960-352-2223	Sequence 2223, Ap
C 34	53	7.4	15720	9	US-10-025-380-1058	Sequence 1058, Ap
C 35	53	7.4	15720	10	US-09-922-217-1058	Sequence 1058, Ap
C 36	53	7.4	15720	10	US-09-833-263-1058	Sequence 1058, Ap
C 37	51.8	7.2	479	10	US-09-864-761-768	Sequence 768, Ap
C 38	51.8	7.2	830	10	US-09-864-761-19531	Sequence 19531, A
C 39	51.4	7.2	396	9	US-09-970-966-53	Sequence 53, Appl
C 40	51.4	7.2	396	10	US-09-825-294-53	Sequence 53, Appl
C 41	51.4	7.2	462	9	US-10-083-357-341	Sequence 341, App
C 42	51.2	7.2	463	10	US-09-960-352-7186	Sequence 7186, Ap
C 43	51	7.1	600	10	US-09-864-761-6690	Sequence 6690, Ap
C 44	50.8	7.1	439	10	US-09-960-352-1009	Sequence 1009, Ap
C 45	50.8	7.1	516	10	US-09-960-352-5785	Sequence 5785, Ap

## ALIGNMENTS

RESULT 1  
US-09-847-173-1  
Sequence 1, Application US/09847173  
Publication No. US20020182228A1  
GENERAL INFORMATION:  
APPLICANT: Collins, Peter J.  
TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94111-3634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/847,173  
CLASSIFICATION: <Unknown>  
FILING DATE: 03-May-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/720,132  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 15280-250-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15223 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other

Thu May 1 14:50:41 2003

us-09-462-816-3.rnpb

Page 2

```

;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

```

Query Match	91.68;	Score 654.8;	DB 9;	Length 15223;
Best Local Similarity	94.88;	Pred. No. 1.4e-164;		
Matches 677; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;

QY	1	CACAAATCAGCTAACCACTGCATATCATACAGATGCAACAGGAGTCAAGAACACA	60
Db	4887	CACAAATCAGCTAACCACTGCATATCATACAGATGCAACAGGAGTCAAGAACACA	4946
QY	61	ACCCCAACATCTCTACTAGATCTCTAGCTTGGATCAGCTTCTCCATCTGTCTGA	120
Db	4947	ACCCCAACATCTCTACTAGATCTCTAGCTTGGATCAGCTTCTCCATCTGTCTGA	5006
QY	121	ATTACATCACAACACACACCTACTACTAGCTTCACAACAACGAGAGTCAAGTCAAACTG	180
Db	5007	ATTACATCACAACATCAGCCATCTACTAGCTTCACAACAACGAGAGTCAAGTCAAACTG	5066
QY	181	CAACCCCAACAGTCAAGACTAAATAAACACAAACAACCCCAACAACCCAGCAAGCC	240
Db	5067	CAATCCACAACAGTCAAGACTAAATAAACACAAACAACCTCAACAACAACCCAGCAAGCC	5128
QY	241	ACTACAAAACAGGCGCAAAACCAACCAACCAAAACCAATATATGATTTTCACTTCGA	300
Db	5127	ACCACAAAACAGGCGCAAAACCAACCAACCAAAACCAATATATGATTTTCACTTCGA	5188
QY	301	GTGTTTAACTTTTACCTTGACAGCATATGCAGCAACAATCCAACTGGTGAGCTATCTGC	360
Db	5187	GTGTTCAACTTTTGAACCTTGACAGCATATGCAGCAACAATCCAACTGGTGAGCTATCTGC	5248
QY	361	AAAAGATTACCAAAACAAAAACCGAGAAAGAAAAACCAACCAAGCTTTCAAAAAACCA	420
Db	5247	AAAAGATTACCAAAACAAAAACCGAGAAAGAAAAACCAATACCAACCCCAAAAAAACCA	5306
QY	421	ACCTTCAGAGCAACCAAAAAAGATCTCAAAACCTCAACACATCACTTAAACCAAGGAATACCC	480
Db	5307	ACCTTCAGAGCAACCAAAAAAGATCTCAAAACCTCAACACATCACTTAAAGGAATACCC	5366
QY	481	ACCAACAGCCACAGAGAGGCCAACCATTCACACCAACCAACCAATCAACATCTCA	540
Db	5367	ACCAACAGCCACAGAGAGGCCAACCATTCACACCAACCAACCAATCAATCTCA	5428
QY	541	CTGCTCACCAACAACACAGAGGAATTCAAAACTCACAAGTCAANTGGAAACCTTCCAC	600
Db	5427	CTACTACCTCTCAACACACAGGAATTCAGAACTCACAAGTCAANTGGAAACCTTCCAC	5486
QY	601	TCAACCTCTCGAAGGCAATCTAAGCCCTTCAAGTCTGCACAACATCGAGACCCA	660
Db	5487	TCAACTCTCTCGAAGGCAATCTAAGCCCTTCAAGTCTCTACAAACATCGAGTACCCA	5548
QY	661	TCAACAACCCATCTCCACCAACACAAACAGCCAGTATCTTAAAAAATAAAAAA 714	
Db	5547	TCAACAACCTTATCTCCACCAACACAAACAGCCAGTATCTTAAAAAATAAAAAA 5600	

RESULT 2  
US-09-844-645-4

```

; GENERAL INFORMATION:
; APPLICANT: Bill

```

```

; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; ADMINISTERING PNEUMOCOCCAL DNA
;

```

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESS: Curtis, Morris & Satorf, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

```
;
;      ZIP: 10036
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/844,645
; FILING DATE: 27-Apr-2001
; CLASSIFICATION: <Unknown>
;
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,505
; FILING DATE: 04-Dec-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
; -S-09-844-645-4
```

Query Match	62.3%;	Score 445.8;	DB 10;	Length 696;
Best Local Similarity	95.3%;	Pred. No. 2.2e-109;		
Matches 470;	Conservative	0;	Mismatches 22;	Indels 1;
				Gaps 1;

QY	1	CACAAGGTACACTAACAACTGCATATATCAAGATGCAACAAAGCCAAATACAAACACA	60
Db	199	CACAAGGTACACCAACAAGTGCATATATCAAGATGCAACAAAGCCAAATCAAGAACACA	258
QY	61	ACCCCAACATACCTCTACAGGATCTCTACGTTGGAAATCAGCTTCTCCATCTGTCTGAA	120
Db	259	ACCCCAACATACCTCTACCCGAATCTCTACGTTGGAAATCAGCTTCTTAATCTGTGAA	318
QY	121	ATTACATCTCAACACCACCATCTACTAGTTCAACAACACCGAGAGTAAAGTCAAACTCG	180
Db	319	ATTACATCTCAAAATACCATCATCTAGTTCAACAACACCGAGAGTAAAGTCAAACTCG	378
QY	181	CAACCCACACAGTACAGTAAAGTCTAAAAACACACAAAC-ACCCCAACACAAACCCAGACGCC	238
Db	379	CAATTCACACAGTACAAACCCAAAAACACAAACAACTCAAAACACAAACCCAGACGCC	438
QY	240	CATTACAAAACAAGCCAAAACCAAAACCAACCAACCAACCAATTAATGATTTTCACTTGA	298
Db	439	CACACAAAACAAGCCAAAACCAAAACCAACCAACCAACCAATTAATGATTTTCACTTGA	498
QY	300	AGTGTTTAACTTTGTACCCCTGCAGCATATGTCAGCAACAATCCAACTGCTGGGCTATCTG	358
Db	499	AGTGTTCAACTTTGTATCCCTCTCAGCATATGTCAGCAACAATCCAACTGCTGGGCTATCTG	558
QY	360	CAAAAATAATACCAAAACAAAACACGAGAAAGAAACACACACCAAGCCATACAAAAAAC	419
Db	559	CAAAAATAATACCAAAACAAAACACGAGAAAGAAACACCTACTACCAAGCCACAAAAAAC	618
QY	420	AACCTTCAAGACACCAAAAAAGATCTCAAACTCAAAACCTAAACCAAGGAATAC	479
Db	619	AACCTTCAAGACACCAAAAAAGATCTCAAACTCAAAACCTAAATATCAAGGAATAC	787
QY	480	CACACCAAGGCC 492	
Db	679	CACACCAAGGCC 691	

RESULT 3  
US-09-878-574-4304/c  
; Sequence 4304, Application US/09878574



```

; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4304
; LENGTH: 793
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(793)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-C11
US-09-878-574-4304

```

```

Query Match          11.7%; Score 83.6; DB 10; Length 793;
Best Local Similarity 48.4%; Pred. No. 1.2e-12;
Matches 276; Conservative 0; Mismatches 292; Indels 2; Gaps 2;

```

```

QY 1 CACAAAGTCACACTACCACTGCAATCTACAGATGCAACAGCCAGATCAAGACACA 60
DB 700 CACCAAAANACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 641
QY 61 ACCCCA-ACATACCTCAGTCCAGATCCAGATGAGTTCAGCTTCCATCTCTCTGA 119
DB 640 AACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 581
QY 120 AATTACATACCAAC-CACCAACCATACTAGCTTCAACAACACAGAGTCAAGTCAAC 178
DB 580 AACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 521
QY 179 TGCAACCCCAACAGTCAAGCTTAACCAACCAACCAACCAACCAACCAACCAAC 238
DB 520 CACACACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCA 461
QY 239 CCACCTACCAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAAC 298
DB 460 ACACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 401
QY 299 AAGTCTTAACCTTTGTACCTGACGATATGACGACCAATTCACCTGCTGGCTATCT 358
DB 400 ACAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 341
QY 359 GCAAAAGAAATGCAAAACCAAAACCAAGAAAGAAACCAACCAACCAACCAAC 418
DB 340 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 281
QY 419 CAACCTTCAAGACACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGTAC 478
DB 280 AAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 221
QY 479 CCACACACCAAGCCACAGACAGACCAACCAATCAACCAACCAACCAACCAACTA 538
DB 220 CCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 161
QY 539 CACTGCTCACCACACACCAACCAACCAAGAAATC 568
DB 160 CACCAACCAACCAACCAACCAACCAACCAAC 131

```

```

RESULT 4
US-09-878-574-4296/C
; Sequence 4296, Application US/09878574
; Patent No. US20020110548A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4296
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(576)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-013-Q1-B1-G1
US-09-878-574-4296

```

```

Query Match          10.7%; Score 76.6; DB 10; Length 576;
Best Local Similarity 50.9%; Pred. No. 7.4e-11;
Matches 207; Conservative 0; Mismatches 199; Indels 1; Gaps 1;

```

```

QY 186 CACACAGTCAAGACTAATAAACACACACACCAACCAACCAACCAACCAACCA 244
DB 556 CACAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 497
QY 245 CAAACCAACCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 304
DB 496 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 437
QY 305 TTAACCTTGTACCTGACGATATGACGACCAATCCATCTGCTGATCTGCAAA 364
DB 436 CCAAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 377
QY 365 GAATGCCAAACCAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCA 424
DB 376 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 317
QY 425 TCAAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAAAGAAAGTACCC 484
DB 316 ACACACACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 257
QY 485 CCAAGCCCAAGAAAGAGCAACATCAACACCAACCAACCAACCAACCAACCA 544
DB 256 CAAAAACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 197
QY 545 TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 591
DB 196 CCACCCACCCCAAAACCAACCAACCAACCAACCAACCAACCAACCA 150

```

```

RESULT 5
US-09-983-965-2109/C
; Sequence 2109, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678

```





```
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 19241
;; LENGTH: 1075
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL078472.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
;; OTHER INFORMATION: EST_HUMAN HIT: AV739739.1, EVALUATE 1.00e+00
;; OTHER INFORMATION: NT HIT: AL163201.2, EVALUATE 2.00e-19
US-09-864-761-19241
```

```
Query Match          10.0%: Score 71.4; DB 10; Length 1075;
Best Local Similarity 45.7%: Pred. No. 2.4e-09;
Matches 249; Conservative 0; Mismatches 296; Indels 0; Gaps 0;
```

```
QY 16 ACAACTGCAATCATACAGATGCAACAAGCCAGATCAAGACACACCAACCCCAATACCTC 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 711 ACCACACCATCATACACACACACACACACACACACACACACACACACACACACACAC 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 ACTCAGAGATCCAGCTTGAGATGAGCTTCCAGCTTCTGCAATATACATCAACAAC 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 651 ACTACACACACACACACACACACACACACACACACACACACACACACACACACACAC 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 ACCACCATCATGCTTCAACAACACACAGAGATCAAGTCAAACTTCAACCCACACAGTGC 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 ACCACCATCATGCTTCAACAACACACAGAGATCAAGTCAAACTTCAACCCACACAGTGC 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 AAGACTAAAAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 531 ACCACCATCATGCTTCAACAACACACAGAGATCAAGTCAAACTTCAACCCACACAGTGC 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 CAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 ACTACCATCATGCTTCAACAACACACAGAGATCAAGTCAAACTTCAACCCACACAGTGC 412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 CCCTGAGCATATGCAACAACATCCACCTGCTGGGCTATCTGCAAAAGATACCAAC 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 ACCAGCACCACACACACACACACACACACACACACACACACACACACACACACACAC 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 AAAAAACCAAGAAAGAAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 ATCACCACACACACACACACACACACACACACACACACACACACACACACACACAC 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 AAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACT 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 ATCACCACACACACACACACACACACACACACACACACACACACACACACACACAC 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 ACCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 ACCAC 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 ACCAC 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 10
US-09-864-761-2513/c
```

```
;; Sequence 2513, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
```

```
;; FILE REFERENCE: Aeonica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 2513
;; LENGTH: 1403
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL078472.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
US-09-864-761-2513
```

```
Query Match          10.0%: Score 71.4; DB 10; Length 1403;
Best Local Similarity 45.7%: Pred. No. 2.7e-09;
Matches 249; Conservative 0; Mismatches 296; Indels 0; Gaps 0;
```

```
QY 16 ACAACTGCAATCATACAGATGCAACAAGCCAGATCAAGACACACCAACCCCAATACCTC 75
```

Db	1061	ACCACACCATATACACCCACACCCATCATCCACACCAACCATATACAC	1002
QY	76	ACTCAGGATCTCTACGTTGGATTCAGCTTCTCCAAATCTGTGTAATTTACATCAACAACC	135
Db	1001	ACTACACACCCACACACACCATCATCCACACCATCCACCATCATCCACACCATCATCCAC	942
QY	136	ACCACCATCTAGTACGTTCAACCAACACGAGSAGTCAAGTCAAACTGCAACCCACACAGCTC	195
Db	941	ACCACCATATACCACTATCAACACACACACACACACATCAGCATCAACACACTACACACACAC	882
QY	196	AAAGCTAAAAACACAACAACAACCCCAACACAACCCAGCAAGCCCACTACAAAAACAACGC	255
Db	881	ACCACACCATCATCACTCAACACACACACACACACATCAGCATCAACACACATCATCAGCATCACC	822
QY	256	CAAAACAAACACCAACAACCAACCAATTAATGATTTTTCACCTTCAAGTGTTTAACCTTTGTA	315
Db	821	ACTACCATCTACACACACACACACACCATCATCAGCATCAACACACACACATCAGCATCACC	762
QY	316	CCCTCAGCATCTGCAAGCAACATCCAACTGCTGGGCTATCTGCAAAAAGATATCCAAAC	375
Db	761	ACCACACACACACACATCAACACACACACACACACATCAAGTACACATCAACACATCAGCATCACC	702
QY	376	AAAAAACCAAGCAAGAAACCAACACACACATCAAGCTTACAAAAAACCAACCTTCAAGSACACC	435
Db	701	ATCACACACACACACATCAACATCAACATCAACATCAACACACACACATCAACATCAAC	642
QY	436	AAAAAAGATCTCAAAACCTCAACCACTAAACCAAGSAGTACCACACACACACACAC	495
Db	641	ATCAGCAGCATCAACACACACACACATCAACATCAACACACACACACATCAACATCAAC	582
QY	496	GAAGACCAACCATCAACACACACCAACAACATCAACAATCAACATCAACATCAACATCAAC	555
Db	581	ACCATCAGCAGCATCAACACACACACACACATCAACATCAACATCAACATCAACATCAAC	522
QY	556	ACCAC 560	
Db	521	ACCAC 517	

```

? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
? SEQ ID NO 20241
? LENGTH: 1635
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC006547.9
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
? OTHER INFORMATION: EXPRESSED IN HEPLA, SIGNAL = 8.6
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-20241

Query Match 10.0%; Score 71.2; DB 10; Length 1635;
Best Local Similarity 46.6%; Pred. No. 3.3e-09;
Matches 261; Conservative 0; Mismatches 298; Indels 1; Gaps 1;

QY 133 ACCACCACCATCTAGTTCCTTCAACCAACACACAGAGTCAGTCAAACTGTGCAACCCACACACA 192
Db 794 ACCACCACCATCTACACACACACACACATCACTACACACACACACACACACACACATCACT 853
QY 193 GTCAAGACTTAAATAAACAACAACAACCAACACACACACACACACACACACACACACACAA 252
Db 854 ACCACCACACACACACACATCACTACACACACACACACACACACACACACACACACACAC 913
QY 253 CGCCAAACCAACCAACCAACCAACCCCAATATGATTTTTCACCTTGGAAAGTTTAACTTT 312
Db 914 ACCACCACCATCTACACACACACACACACCACTATCACTACACACATCAACACACACACAC 973
QY 313 GTACCTCTGACGATATGTGAGCAACATCACTGCTGGGTATCTATGTGAAAAGATATACA 372
Db 974 ACCATCACACACACACACATCACTACATCTACACACACACACATCTACACACACACACAC 1033
QY 373 AACCAAAAAACGAGAAAGAAACACACACCAAGCCTTACAAAACCAACCAACCTTCAAGACA 432
Db 1034 ATCACTTACACACACACACACACACACCACTACTACTACACACACACACACACACATCA 1093
QY 433 ACCCAAAAAAGATCTCAAACTTCAAACCAACCACTAAACCAAGAAAGTAATCCACACACAC 492
Db 1094 ACCCAACACACACACACATCTACACACACACACACACACACACATCTACACACACACAC 1153
QY 493 ACAGAAAGCCACACATCAACACACACCAACCAACCAATCAACATCTGCTCAACAC 552
Db 1154 ATCAACATCACTACACACATCAACACACACACACACACACACACATCAACACACACATCA 1213
QY 553 AACACACAGGAATCCAAAACTCACAAGTCAATGGAAACCTTGCACCTCAACCTCTCC 612
Db 1214 ACTACACACACACACACACATCACTACACACACATCAACACACATCAACACATCAACACAC 1273

```

OY	613	GAAGGCAATTAAGCCCTTCTCAAGTCTCCACA	-CATCCGAGCAGCCATCAAAACCCCTC	671
Db	1274	ACAACCAACATCACTACTACGACATCACCACACTTACCACCATCTACTACCAACCAACCAACCAAC		1353
OY	672	ATCTCCACCCACACAGAC		691
Db	1334	ATCAACCATCACTTCCACAC		1353

```

RESULT 12
US-09-864-761-3471
; Sequence 3471, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;   APPLICANT: Penn, Sharon G.
;   APPLICANT: Rank, David R.
;   APPLICANT: Hanzel, David K.
;   APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3471
; LENGTH: 1973
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006547.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
; OTHER INFORMATION: EXPRESSED IN HBL10C, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11

```

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8-4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-3471

```

Query Match	10.0%;	Score 71.2;	DB 10;	Length 1973;
Best Local Similarity	46.6%;	Pred. No. 3.6e-09;		
Matches 261;	Conservative	0;	Mismatches 298;	Indels 1;
				Gaps 1

OY	133	ACACGACGACACTAGTGGTTGACACACGAGGTGAAAGTCAAACTGGAAACCCGACACA	192
Db	948	ACCCACACCATCACCACACACACACACCATCAGTACCTACCAACCAACCAACCACTACCT	1007
OY	193	GTCAGAGCTAAAAACACACACACACCCCAACCAACCCGAGACGCCACTGCAAAACAA	252
Db	1008	ACCAACACACACACACACATCAGTACGACACACACCATCATGACACACGACAC	1067
-OY	253	CGCCAAAACAAACCCAAACAAACCCCATATGATTTTCTTCTTGAAAGTTTAACTTT	312
Db	1068	ACCCACATCAGTACACACACACACACATCAGCATCTACCAACCATCAGCACAC	1127
OY	313	GTACCTCGACGACATATGCAGCAACATCAACCTCGTGGGTATCGAAAAAGATTACA	372
Db	1128	ACCATCACCACACACCATCATCATCTAGTACCAACCAACATATCAACCAACACAC	1187
OY	373	AACAAAAAACGAGAAAGAAACACACACCAAGCCTACAAAAAACCAACTTTCAGACA	432
Db	1188	ATCACTACACACACACACACACACACATCATCTACACACACACACACACATCAT	1247
OY	433	ACCAAAAAAGATCTCAAACTCAAAACCTAAACCAAGAAAGTACCCACACCAAGCC	492
Db	1248	ACCCACACACACACATCAGTACACACACACACACACCATCATCAACCAACCAAC	1307
OY	493	ACAGAAAGGCAACATCAACACACACAAACCAACATCAACATCACTGGTGCACAC	552
Db	1308	ATCAACATCACTACACCATCACACACACACACACATCAACCAACCAACATCAATC	1367
OY	553	AACACACAGGAAATCCAAAACTCACAGTCAATGGAACCTTCAGTCAACCTCTCC	612
Db	1368	ACTGACACACACACACACATCATCTACCAACCATCAACACCATTAACGACCATGACT	1427
OY	613	GAAAGCAATCTAAGCCCTTCTCAAGTCTCCAAA-CATCGAGCAACCATCAACACCTC	671
Db	1428	ACAACACACATCACTACACCATACACATACACATACCAACCATCATCAACCAACAC	1487
OY	672	ATCTTCACCCACACACAC 691	
Db	1488	ATCAACATCACTTCCACAC 1507	

```

RESULT 13
US-09-864-761-26582/c
; Sequence 26582, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04

```







GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:08:32 ; Search time 884.239 Seconds  
(without alignments)  
13095.757 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaaagtcacactaacac.....glagtattcaaaaaaaaaa 715

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 16154066 segs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estoy:\*  
6: em\_estpl:\*  
7: em\_estio:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pla:\*  
21: em\_gss\_vre:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	120	16.8	895	17	CNS0071A
C 2	104.6	14.6	946	17	AG127412
C 3	103.2	14.4	922	17	CNS0073W
C 4	101.8	14.2	919	17	CNS04EMX
C 5	100	14.0	1201	9	AL581589
C 6	99	13.8	1101	17	CNS00EXE

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CNS0071A/C	CNS0071A	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR14809 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL066286	AL066286.1	GI:4945153	GSS.	Drosophila melanogaster.	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	1 (bases 1 to 895)			
C 7	95.8	13.4	1097	17	CNS01224							AL102202 Drosophila
C 8	95	13.3	791	17	CNS0009KS							AL053801 Drosophila
C 9	95	13.3	1101	17	CNS0010LO							AL060607 Drosophila
C 10	94.4	13.2	1216	17	AG135357							AG135357 Pan trogl
C 11	94.2	13.2	815	17	B12686							B12686 F27J14-T7 I
C 12	93.8	13.1	1008	17	AG137085							-5137085 Pan trogl
C 13	93.6	13.1	884	17	CNS0060U							AL063923 Drosophila
C 14	93.6	13.1	969	17	AG0743309							AG0743309 HS_5387_B
C 15	93.2	13.0	1141	17	AG0743305							AG0743305 HS_5387_B
C 16	92.8	13.0	981	17	AG127518							AG127518 Pan trogl
C 17	92.6	13.0	1183	17	AG136828							AG136828 Pan trogl
C 18	92.4	12.9	859	17	AG128925							AG128925 Pan trogl
C 19	92.4	12.9	880	17	AG139490							AG139490 Pan trogl
C 20	92	12.9	1137	17	AG0743326							AG0743326 HS_5387_B
C 21	92	12.9	1225	17	CNS0166K							AL106358 Drosophila
C 22	91.4	12.8	700	17	AG127423							AG127423 Pan trogl
C 23	91.4	12.8	885	17	CNS0317M							AL249907 Tetrarodon
C 24	91	12.7	1099	17	AG0743360							AG0743360 HS_5387_B
C 25	90.8	12.7	1081	17	AG135328							AG135328 Pan trogl
C 26	90.6	12.7	1491	12	BE882936							BE882936 601505836
C 27	90.2	12.6	866	17	AG126308							AG126308 Pan trogl
C 28	90	12.6	953	17	CNS006PF							AL063557 Drosophila
C 29	90	12.6	1101	17	CNS017FC							AL107970 Drosophila
C 30	89.8	12.6	896	17	AG141027							AG141027 Pan trogl
C 31	89.4	12.5	909	17	CNS00JTL							AL076720 Drosophila
C 32	89.4	12.5	955	17	AG076494							AG076494 Pan trogl
C 33	89.4	12.5	1024	17	AG133080							AG133080 Pan trogl
C 34	89.4	12.5	1054	17	AG137135							AG137135 Pan trogl
C 35	89.4	12.5	1188	17	AG135332							AG135332 Pan trogl
C 36	89.4	12.5	1353	17	AG128010							AG128010 Pan trogl
C 37	89	12.4	1101	17	CNS00LT2							AL076720 Drosophila
C 38	89	12.4	1144	17	AG0743364							AG0743364 HS_5387_B
C 39	88.8	12.4	938	17	AG132199							AG132199 Pan trogl
C 40	88.4	12.4	926	17	AG078232							AG078232 HS_3176_B
C 41	88.2	12.3	966	17	AG126022							AG126022 Pan trogl
C 42	88.2	12.3	987	17	CNS00K3Q							AL077186 Drosophila
C 43	88.2	12.3	1101	17	CNS010B3							AL098745 Drosophila
C 44	88	12.3	860	17	CNS018FL							AL109275 Drosophila
C 45	88	12.3	960	17	CNS005F3							AL059925 Drosophila

## ALIGNMENTS

RESULT 1  
CNS0071A/C  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

895 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR14809 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
AL066286  
AL066286.1 GI:4945153  
GSS.  
Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 895)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
determination of this BAC end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Otsuoka and  
Aaron Mammosser in Pletier de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial



OY		615	AAGCAATCTAACCCCTTGTCAAGTCGACAAACAATCGCAGACCATTACAAACCTCATC	674
Dd		753	CAGCACCCCATAAAAAAAAAAACCCTCCACGAACAAAAAAAAGAAGACAAACAACGCCACAA	812.
OY		675	TGCACCACACACACGGCCAGTAGTTATTAATAAAAAAAAAA	714
Dd		813	CCAANAACCCACAGAAAAACACACACGCCCCAACACAAAAA	852
<hr/>				
RESULT 3 CONS0073M/c				
LOCUS DEFINITION      CNS0073M                  922 bp       DNA           linear    GSS-03-JUN-1999				
DEFINITION      Drosophila melanogaster genome survey sequence TET3 end of BAC #				
ACCESSION      AL066784				
VERSION      AT066784.1 GI:4945247				
KEYWORDS      GSS.				
SOURCE			Drosophila melanogaster.	
ORGANISM			Drosophila melanogaster.	
JOURNAL			Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
			Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
			Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE			1 (bases 1 to 922)	
AUTHORS			Genoscope.	
TITLE			Direct Submission	
COMMENT			Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :	
			BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
			- Web : www.genoscope.cns.fr )	
			determination of this BAC-end sequence was carried out as part of a	
			collaboration with The Berkeley Drosophila Genome Project (BDGP).	
			The BDGP is constructing a physical map of the Drosophila	
			melanogaster genome using these BACS. For further information	
			Please see http://www.fruitfly.org The BDGP Drosophila	
			Melanogaster BAC library was prepared by Kazutoyo Osoegawa and	
			Aaron Mamooser in Pieter de Jong's laboratory in The Department of	
			Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,	
			N.Y. The library is named RPcI-98 and was constructed by partial	
			EcoRI digestion of Drosophila DNA provided by the BDGP from the	
			Isogenic strain Y2; cn bw sp, the same strain used for the BDGP's	
			P1 and EST libraries. A more detailed description of the library	
			and how to order individual BAC clones, the entire library, or	
			filters for hybridization from the BACPAC Resource Center can be	
			found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
FEATURES			Location/Oualifiers	
Source			1..922	
			/organism="Drosophila melanogaster"	
			/db_xref="taxon:7227"	
			/clone="BACR1AD09"	
			/clone_lib="RPcI-98"	
			/note="end : TET3"	
BASE COUNT			223 a     95 c     109 g     221 t     274 others	
ORIGIN				
			Query Match                  14.4% ; Score 103.2 ; DB 17 ; Length 922 ;	
			Best local similarity        24.8% ; Pred. No. 5.2e+10 ;	
			Matches 102 ; conservative 164 ; mismatches 144 ; Indels 1 ; Gaps 1 ;	
OY		181	CAACCCAGACACTCAAGACTATAAAACAACACACACCCCAAACACCAACCCGACAAGCC	240
Dd		921	NNNNMMCMCCMCOCCKKMMNKKMKNMCMKMKNMNMMNKMMNMMCNMMNMMNMMNMAC	862
OY		241	ACTACAAAACAAGCCAAACAACACCAACCAACCAATATGATTTTGACTTCGAA	300
Dd		861	ACAMMAANNMMNAKCMMMNCMYAKCMKMMACMMACMMMAAMMMMMMMNMMMACAMMAN	802*
OY		301	GTTTTAAGTTTGATCCCGGACGAFATGACACCAACCAATCAACCTGGTGGGCTATNGC	360
Dd		801	ACACGACAMMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMNMMN	743*
OY		361	AAAAGATACCAACAAAAAACCAAGAAAAGAAACACACACACCAACCTACAAAAAACCA	420

[illegible]

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see: <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammoser in Pierre de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial *NotI* digestion of *Drosophila* DNA provided by the BDGP from the







```

source 1. 815
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db.xref="taxon:3702"
/clone="F27J14"
/sex="hermaphrodite"
/notes="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT 22 a 11 c 185 g 302 t 95 others
ORIGIN

Query Match 13.2%; Score 94.2; DB 17; Length 815;
Best local similarity 45.0%; Pred. No. 2.6e-08;
Matches 297; Conservative 0; Mismatches 355; Indels 8; Gaps 1;

QY 15 AACACTGCATCATACAGATGACAGCCAGATCAAGAACACACACCCTCAATACCT 74
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 812 AAAAAAACCAAAACAAAAAACCAACAAAAAACCAAAACAAAAAACCAAAACAAA 753

QY 75 CACTCAGATCCCTCAGCTTGATGAGCTTCCATCTGTCTGAATTCATCAACAAC 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 752 AAACCAAAACCAAAACAAAAAACCAACAAAAAACCAAAACAAAAAACCAAA 693

QY 135 CACCACCATCTAGCTTCAACACACAGAGTCAAGTCAAACTGCAACCCCAACAGT 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 692 AAACACAAACCAAAACAAAAAACCAACAAAAAACCAAAACAAAAAACCAAA 633

QY 195 CAAGACTAAAAACACACACACCAACACACACCCAGACCCCACTACAAACACAG 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 632 AAACACAAAAAACCAACACACACCCCAAAACCAACAAAAAACCAACAAACACA 573

QY 255 CCAAAACAAACCAACCAACCAACCAATATGATTTTCACTTGAAGTTTAACCTTGT 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 572 AAAAAAACCAAAACCAAAACAAAAAACCAAAACAAAAAACCAAAACAAAA 513

QY 315 ACCCTGACATATGACAGCAACATCAACCTGCTGGCTATCTGCAAAAGATACCAA 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 512 AAAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 453

QY 375 CAAAAAACCGAAGAAACCAACCAACCAAGCTTACAAAAAACCAACCTTCAAGCAAC 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 AAAAAAACCAAAACCAAAACCAACCCCNCCNCCNCCNCCNCCNCCNCCNCCN 393

QY 435 CAAAAAGATCTCAACCTCAACCA-----CTAAACCAAGAGAGTACCAACAC 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 NNACACAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 333

QY 487 AAGCCACAGACAGCCACATCAACACCAACCAACCAACCACTCACTACTGCTC 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 AAACCAACCAAAACCAACCAACCAAAACCAAAACCAAAACCAACCAACCA 273

QY 547 ACCAACACACACAGGAATTCAAAACCTCAACAGTCAATGGAACCTTCCACTCAAC 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 ACAAAAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 213

QY 607 TCCTCCAGAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCATCA 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 CCACACACACCAAAACCAAAACCAAAACCAACCCCAACCCCAACCAAA 153

RESULT 12
AG137085 1008 bp DNA linear GSS 04-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-150N19.F, genomic survey sequence.
DEFINITION AG137085
ACCESSION AG137085
VERSION AG137085.1 GI:16666763
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone_11b:PTB Chimpanzee Male
          BAC library clone:PTB-150N19.F.
          Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

```

```

REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
         Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1008)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
         Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
         and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
         1-7-22 Suehiro-cho,Tsukuba, Ibaraki, Japan
         (E-mail:chimbases@gsr.riken.go.jp, URL:http://hpg.gsr.riken.go.jp/,
         Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
         was generated during the R&D process and may have higher chance of
         clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pRS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 1008
/organism="Pan troglodytes"
/db.xref="taxon:9598"
/clone="PTB-150N19.F"
/sex="male"
/cell_type="lymphoblast"
/clone_11b="PTB Chimpanzee Male BAC library"
BASE COUNT 394 a 498 c 41 g 29 t 46 others
ORIGIN

Query Match 13.1%; Score 93.8; DB 17; Length 1008;
Best local similarity 47.8%; Pred. No. 3.1e-08;
Matches 272; Conservative 0; Mismatches 297; Indels 0; Gaps 0;

QY 119 AATTATCATCAACCAACCAACCACTACTAGCTTCAACACACAGAGTCAAGTCAAAC 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 ACACACACCCCAAAACCAACCCCAACCAACCAAAACCAAAATTTCTGCACTATCC 485

QY 179 TGCACCCCAACAGTCAAGACTTAAAAACACACACCAACCAACCAACCCGACAGC 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 CCCCACCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 545

QY 239 CCATACCAAAACGCAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTCG 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 CAAAAAACCAACCAACCAACCAACCCCGCACACCCCAACCAACCAACCAACCA 605

QY 299 AAGTGTATTAAGTGTACCTGACGATATGACGACATCAATCAACCTGCTGGCTATCT 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 CAGAACCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 665

QY 359 GCAAAAGATACCAACCAACCAACCAAGAAAGAAACCAACCAACCAAGCTTCAAAAA 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 725

QY 419 CAACCTTCAAGACCAACCAAAAGATCTCAACCTCAACCACTTAAACCAAGAGATAC 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 726 CCGACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 785

QY 479 CCACACCAAGCCCAAGAGACCAACCAACCAACCAACCAACCAACCAACCAACCA 538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 786 CACACACCAACCAACCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCA 845

QY 539 CACTGCTCAACCAACCAACCAAGGAATCCAAAACCTCAAGTCAATGGAATGCACTCC 598
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 846 AACACCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 905

QY 599 ACTCAACCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGC 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 906 ACAAAACACACCAACCAACCAACCCCAACCAACCCCAACCAACCAACCAACCA 965

```



QY	659	CATCACACCCCTACTCCACCACACA	687
Db	966	CACCAACACCAAAACCACACACAAA	994
RESULT 13			
CNS006U0/c			
LOCUS			
DEFINITION		CNS006U0 884 bp DNA linear GSS 03-JUN-1999	
ACCESSION		Drosophila melanogaster genome survey sequence T7 end of BAC #	
VERSION		BACFI4N21 of RPcI-98 library from Drosophila melanogaster (fruit	
SOURCE		fly), genomic survey sequence.	
ORGANISM		AL065923	
KEYWORDS		GSS.	
REFERENCE		Drosophila melanogaster.	
AUTHORS		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
TITLE		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
JOURNAL		Ephyroidae; Drosophilidae; Drosophila.	
COMMENT		Genoscope. 1 (bases 1 to 884)  Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequence; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammossier at Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPcI-98 and was constructed by partial EcoRI digestion of cn bw sp. DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bdacpac.med.buffalo.edu/drosophila_bac.htm.  Location/Qualifiers 1..884 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACFI4N21" /clone_lib="RPcI-98" /note="end : 77"	
FEATURES			
SOURCE			
BASE COUNT		230 a 62 c 139 g 124 t 329 others	
ORIGIN			
Query Match		13.1% Score 93.6; DB 17; Length 884;	
Best Local Similarity		17.4%; Pred. No. 3.4e-08;	
Matches		57; Conservative 167; Mismatches 103; Indels 0; Gaps 0;	
QY	369	ACCAAACAAAAACAGGAAAGAAGAACACACACCAAGCTACAAAAAACAACCTTCAA	428
Db	884	MNNMNNNNMMHTTKKRTTHPVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	825
QY	429	GACAAACAAAAAATCTGCAAACCTCAAAACCATTAACCAAGGAATACCCACACCAA	488
Db	824	AMNN	765
QY	489	GGCCACAGAGGCCAATCAACACACCAAAAAACAACATGCACACTACACTGTTCAC	548
Db	764	NN	705
QY	549	CAACAACACACAGGAATATCCAAATCTCAATGATCAATGAATCGTTCCACTTCAAC	608
Db	704	CSNNNAACCCCNACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	645
QY	609	CTCCGAGGCAATCTAAGCTCTTCAAGTGCTCCACCAATCGAGAGCCATCATCAAC	668

Db	644	MCCCCCCCMMMAACAACAMCAAMAMCMAMAMAMAMAMAMAMMCCACCACACMC	585
OY	669	CTCATGTCCACCCAACACAAACAGCGCA	695
Db	584	MCCAMCTCACCAACCAAHMMAAAMA	558
RESULT 14			
LOCUS	A0743309	969 bp DNA linear	GSS 16-JUL-1995
DEFINITION	HS_5387_B2_B07_SPC1-11 Human Male BAC library Homo sapiens genomic clone Plate=963 Col=14 Row=D, DNA sequence.		
ACCESSION	A0743309		
VERSION	A0743309.1	GI:5520831	
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 969) Mahairas,G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)		
MEDLINE	99380589		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ( <a href="http://bacpac.med.buffalo.edu/ordering_bac.htm">http://bacpac.med.buffalo.edu/ordering_bac.htm</a> ) or from Resear h Genetics ( <a href="http://www.hscg.washington.edu">http://www.hscg.washington.edu</a> ) Plate: 963 row: D column: 14 Seq primer: SP6 Class: BAC ends High quality sequence stop: 969. Location/Qualifiers 1..969 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="Plate=963 Col=14 Row=D" /clone_lib="RPCI-11 Human Male BAC Library" /sex="male" /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"		
BASE COUNT	647 a 294 c 7 g 5 t 16 others		
ORIGIN			
Query Match	13.1%; Score 93.6; DB 17; Length 969;		
Best Local Similarity	48.0%; Pred. No. 3.4e-08;		
Matches 286; Conservative	0; Mismatches 307; Indels 3; Gaps 1;		
OY	119	AAATTATGATTCACAACACCAACCACTACTAGTTTCAACACACCGGAGTCAAGTCAACC	178
Db	56	AACTAAACCAACNNAACTCTAANNAGGCAAAAAACAAANNNAACAAACAAAACCAACA	115
OY	179	TGCACACCCACAACAGTCAAGACTAAAAACACACACACCAACCAACCAACAGCAGC	238
Db	116	CAAAACCAAAACAAACCAACGCCACACACACACACCCCAAACCAACCAACCAACCC	175
OY	239	CCACTACAAACAACGGCAAAACCAACCAACCAACCAACCAACCAATTAATGATTTCACCTTG	298

```

Db 176 ANCCACACACAAACCAACAA---CMAAAACAAACAAACAAACAAACAAAC 232
Qy 299 AAGTGTAACTTTGTACCTCTGAGCATATGACCAACAAATCCAACTGCTGGCTATCT 358
Db 233 AAACACCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 292
Qy 359 GCAAAAGATATCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 418
Db 293 AAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 352
Qy 419 CAACCTTCAAGCAACCAACAAACAAACAAACCTCAACCTCAACCTCAACCAAGAGTAC 478
Db 353 AACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 412
Qy 479 CCACCAACAGCCCAAGAGAGCCCAATCAACCAACCAACCAACCAACCAACCA 538
Db 413 CCCCACAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 472
Qy 539 CACTGCTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTTC 598
Db 473 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 532
Qy 599 ACTCAACCTCTCCAGGCAATCTTAAGCCTCTCAAGTCTCAACCAACCTCGAGACC 658
Db 533 AACACAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 592
Qy 659 CATCAACACCTCTCTGACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
Db 593 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 648

```

```

RESULT 15
AO743305 1141 bp DNA linear GSS 16-JUL-1999
DEFINITION HS_5387_B2_B02_SP6 RPCI-11 Human Male BAC Library Homo sapiens
ACCESSION AO743305
VERSION AO743305.1 GI:5520827
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Library availability, please contact Pieter de Jong
(piet@edej.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end web Server:
http://www.htsc.washington.edu
Plate: 963 row: D column: 4
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 1141.
Location/Qualifiers
1..1141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=963 Col=4 Row=D"

```

```

/clone.lib="RPCI-11 Human Male BAC Library"
/sex="male"
/Note="Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"
BASE COUNT 760 a 301 c 16 g 15 t 49 others
ORIGIN

```

```

Query Match 13.0%; Score 93.2; DB 17; Length 1141;
Best Local Similarity 48.0%; Pred. No. 3.9e-08;
Matches 284; Conservative 0; Mismatches 306; Indels 2; Gaps 1;

```

```

Qy 124 ACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGGATCAACCTGCA 183
Db 243 ACAACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 302
Qy 184 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 243
Db 303 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 362
Qy 244 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 303
Db 363 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 422
Qy 304 TTTAACTTTGACCTGACATATGACCAACCAACCAACCTGCTGGCTATCTGCA 363
Db 423 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 482
Qy 364 AGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 423
Db 483 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 542
Qy 424 TTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 483
Db 543 CACA-ACANCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
Qy 484 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 543
Db 601 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 660
Qy 544 CTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 603
Db 661 ANANAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 720
Qy 604 ACTCTCTCGAAGCAATCTTAAGCCTTCTCAAGTCTCAACCAACCTCGAAGCA 663
Db 721 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780
Qy 664 CAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
Db 781 CACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 832

```

Search completed: May 1, 2003, 12:33:32  
Job time : 890.239 secs



PT containing the protein G sequence, useful in protective vaccines and  
PT to raise antibodies for diagnosis

PS Claim 9, Fig 3; 67pp; English.

XX The respiratory syncytial virus (RSV) G protein can be used in  
CC vaccines by inserting the G protein gene into a non-replicating  
CC vector. The G protein is placed under the control of alternative  
CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G proteins  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods. This truncated G  
CC protein is secreted since it lacks a transmembrane domain.

XX Sequence 232 AA:

Query Match 100.0%; Score 1229; DB 20; Length 232;  
Best Local Similarity 100.0%; Pred. No. 3e-89;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSOITTLASTTPGVKSNL 60  
DB 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSOITTLASTTPGVKSNL 60  
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPNKPNNDHFHEVENFVPCISNNPTCMAIC 120  
DB 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPNKPNNDHFHEVENFVPCISNNPTCMAIC 120  
QY 121 KRIPNKKPGKTTTKPTKPKPTKTKDKLPQTTPKPEVPTTPTEPTINTTKNITTT 180  
DB 121 KRIPNKKPGKTTTKPTKPKPTKTKDKLPQTTPKPEVPTTPTEPTINTTKNITTT 180  
QY 181 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232  
DB 181 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232

RESULT 2

AAR39286  
ID AAR39286 standard; Protein; 298 AA.

AC AAR39286;

DT 13-JAN-1994 (first entry)

DE Respiratory syncytial virus (RSV) G protein.

KW PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.

OS Respiratory syncytial virus.

PN WO9314207-A.

PD 22-JUL-1993.

PF 05-JAN-1993; 93WO-CA00001.

PR 06-JAN-1992; 92GB-0000117.

PA (CONN-) CONNAGHT LAB LTD;

PI Ewasyshyn ME, Klein MH;

DR WPI: 1993-243222/30.  
N-PSDB: AAQ45686.

PT Multimeric hybrid genes and their chimeric proteins - are

PT vaccines against multiple pathogenic infections e.g.  
PT para-influenza virus and respiratory syncytial virus

PS Claim 11, Figure 7A-7D; 80pp; English.

XX A novel multimeric hybrid gene is used as a vaccine. The gene  
CC consists of two gene sequences which are linked and encode antigenic  
CC regions, these two sequences being derived from two different  
CC pathogens (para-influenza virus (PIV) and respiratory syncytial virus  
CC (RSV)). The gene sequences that are particularly used are those  
CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
CC G proteins (AAQ45685, AAQ45686).

XX Sequence 298 AA:

Query Match 100.0%; Score 1229; DB 14; Length 298;  
Best Local Similarity 100.0%; Pred. No. 4e-89;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSOITTLASTTPGVKSNL 60  
DB 67 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSOITTLASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPNKPNNDHFHEVENFVPCISNNPTCMAIC 120  
DB 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPNKPNNDHFHEVENFVPCISNNPTCMAIC 186  
QY 121 KRIPNKKPGKTTTKPTKPKPTKTKDKLPQTTPKPEVPTTPTEPTINTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKPKPTKTKDKLPQTTPKPEVPTTPTEPTINTTKNITTT 246  
QY 181 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232  
DB 247 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 298

RESULT 3

AAW96313  
ID AAW96313 standard; Protein; 298 AA.

AC AAW96313;

DT 28-JUN-1999 (first entry)

DE Membrane bound G protein of respiratory syncytial virus.

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

OS Respiratory syncytial virus.

PN WO9904010-A1.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-CA00697.

PR 18-JUL-1997; 97US-0896442.

PA (CONN-) CONNAGHT LAB LTD.

PI Klein MH, Li X, Sambhara S;

DR WPI: 1999-132254/11.  
N-PSDB: AAX08421.

PT Immunogenic composition for generating antibodies against  
PT respiratory syncytial virus - comprises non-replicating vector  
PT containing the protein G sequence, useful in protective vaccines and  
PT to raise antibodies for diagnosis  
PS Claim 4; Fig 2; 67pp; English.

```
XX CC The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternating
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods.
XX CC
SQ Sequence 298 AA:
Query Match 100.0%; Score 1229; DB 20; Length 298;
Best Local Similarity 100.0%; Pred. No. 4e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HKVTLTATIIQDATSQIKNTPTLYLTQDPOLGISFNSLSRITTSQTTTLASTTPGVKSNL 60
DB 67 HKVTLTATIIQDATSQIKNTPTLYLTQDPOLGISFNSLSRITTSQTTTLASTTPGVKSNL 126
QY 61 QPTTVKTKNTTTTQTOPSKPTTKORONKPPNKNPNDFEVEFNPCISCSNNPTCWAIC 120
DB 127 QPTTVKTKNTTTTQTOPSKPTTKORONKPPNKNPNDFEVEFNPCISCSNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 180
DB 187 KRIPNKKPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 246
QY 181 LTNNTTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPPTTRQ 232
DB 247 LTNNTTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPPTTRQ 298
RESULT 4
AAP70845
ID AAP70845 standard; protein; 298 AA.
XX AC AAP70845;
XX DT 05-APR-1991 (First entry)
XX DE Sequence of human respiratory syncytial virus (HRSV) A2 strain
XX DE G protein.
XX KW Vaccine.
XX OS Human respiratory syncytial virus (HRSV).
XX PN W08704185-A.
XX PD 16-JUL-1987.
XX PF 23-DEC-1986; 86MO-US02756.
XX PR 14-JAN-1986; 86US-0818740.
XX PA (UYNC-) UNIV OF N CAROLINA.
XX PA (WERTZ) WERTZ G W.
XX PA (WERTZ) WERTZ G W.
XX DR WPI; 1987-206300/29.
XX DR N-PSDB; AAN0784.
XX PT Vaccines for human respiratory virus - comprising proteins or
XX PT fragment encoded by a DNA sequence coding for human respiratory
XX PT syncytial virus proteins.
XX PS Disclosure; Chart 13; 57pp; English.
```

```
XX CC A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
XX CC
SQ Sequence 298 AA:
Query Match 91.1%; Score 1120; DB 8; Length 298;
Best Local Similarity 92.7%; Pred. No. 1.6e-80;
Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
QY 1 HKVTLTATIIQDATSQIKNTPTLYLTQDPOLGISFNSLSRITTSQTTTLASTTPGVKSNL 60
DB 67 HKVTLTATIIQDATSQIKNTPTLYLTQDPOLGISFNSLSRITTSQTTTLASTTPGVKSNL 126
QY 61 QPTTVKTKNTTTTQTOPSKPTTKORONKPPNKNPNDFEVEFNPCISCSNNPTCWAIC 120
DB 127 QPTTVKTKNTTTTQTOPSKPTTKORONKPPNKNPNDFEVEFNPCISCSNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 180
DB 187 KRIPNKKPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 246
QY 181 LTNNTTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPPTTRQ 232
DB 247 LTNNTTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPPTTRQ 298
RESULT 5
AAR25302
ID AAR25302 standard; Protein; 298 AA.
XX AC AAR25302;
XX DT 03-MAR-1993 (First entry)
XX DE HSRV glycoprotein G (gpc).
XX DE HSRV glycoprotein G (gpc).
XX KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
XX KW major capsid protein; N.
XX OS Human respiratory syncytial virus strain A2.
XX PN US5149650-A.
XX PD 22-SEP-1992.
XX PF 14-JAN-1986; 86US-0818740.
XX PR 14-JAN-1986; 86US-0818740.
XX PR 13-JUL-1988; 88US-0218737.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PA Collins PL, Wertz GW;
XX PI Collins PL, Wertz GW;
XX DR WPI; 1992-340247/41.
XX DR N-PSDB; AAO29623.
XX PT Vaccines for human respiratory virus - include structural genes
XX PT coding for native structural viral proteins and immunogenic
XX PT fragments
XX PS Disclosure; Page 18; 21pp; English.
XX CC The sequences of mRNA encoding HRSV structural proteins are given in
XX CC AA029622-26. The proteins are F, G, 22K, 9.5K and major capsid
XX CC protein N. The sequences and encoded proteins are useful for
XX CC preparing vaccines against HRSV. The vaccines can be used to confer
XX CC immunity against respiratory tract infections on human subjects.
```

Seq	Sequence	298 AA
Query Match	91.18%	Score 1120; DB 13; Length 298;
Best Local Similarity	92.78%	Pred. No. 1.6e-80;
Matches 215; Conservative	5; Mismatches 12; Indels 0; Gaps 0;	
QY	1 HKVLTPTAIQDATSQIKNTPTPTVLYLQDLPQGLSPNSLSEITSOITVITLILASTTPGVKSNL 60	
Db	67 HKVPTPTAIQDATSQIKNTPTPTVLYLQDLPQGLSPNSLSEITSOITVITLILASTTPGVKSTL 126	
QY	61 QPTTVTKTKNTTQTQPSKPTTKQKQNPKNPKPNNDHFEVFNNVPCSSISNNPTCAIC 120	
Db	127 QSTTVKTKNTTQTQPSKPTTKQKQNPKNPKPNNDHFEVFNNVPCSSISNNPTCAIC 186	
QY	121 KRIPNNKPKGKKTITTKRPFETTKRKDLQPKTQPKKEVPTTKPEETPIINTKTNITTT 180	
Db	187 KRIPNNKPKGKKTITTKRPFETTKRKDLQPKTQPKKEVPTTKPEETPIINTKTNITTT 246	
QY	181 LTNNTTGNKRLTSQMEFTHSTSSSEGNLSPSQVSTTSSHPSSPSSPENTRQ 232	
Db	247 LTNNTTGNKRLTSQMEFTHSTSSSEGNLSPSQVSTTSSHPSSPSSPENTRQ 298	

Accession	Protein	Length	Score	DB	Length
AA047605	standard, Protein; 298 AA.	298	91.1%	1120	298
AA047605			92.7%	Pred. No. 1.0e-80	
11-JUN-1998	(first entry)				
HRSV glycoprotein G.					
HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.					
Human respiratory syncytial virus.					
US5716823-A.					
10-FEB-1998.					
12-MAY-1997; 97US-0854783.					
13-JUL-1988; 88US-0218737.					
14-JAN-1986; 86US-0818740.					
23-DEC-1986; 86MO-US02756.					
11-JUN-1992; 92US-0891171.					
12-MAY-1997; 97US-0854783.					
(PHAA ) PHARMACIA & UPJOHN CO.					
Collins PL, Wertz GW;					
WPI; 1998-144802/13.					
N-PDB; AAV18736.					
Production of human respiratory syncytial virus glyco-protein F or G					
- by culturing eukaryotic host cells transfected with corresponding					
DNA					
Example 1; Columns 27-28; 17pp; English.					
The present sequence was used in the development of a novel method					
for the production of human respiratory syncytial virus (HRSV)					
glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises					
culturing eukaryotic host cells transfected with an isolated DNA					
sequence encoding HRSV gpF or gpG. The gp can be used to prepare					
vaccines against HRSV.					
Sequence 298 AA:					
Query Match					
Best Local Similarity					

Matches	21;	Conservative	5;	Mismatches	12;	Indels	0;	Gaps	0;
QY	1	HKVLTPTAIIOQATSOIKNTPTPTVYLTDLPDQIGISFISNLSSETTSOTTTIIASTPQVKSNI	60						
Db	67	HKVPTPTAIIOQATSOIKNTPTPTVYLTDLPDQIGISFISNLSSETTSOTTTIIASTPQVKSNI	126						
QY	61	QPTPTVKNTNTTQDQPSKPTTKORONRPPKPNNDFFEFVFNVPQCSISNNPTCMALC	120						
Db	127	QSTTVTKTKNTTTQDQPSKPTTKQONRPPKPNNDFFEFVFNVPQCSISNNPTCMALC	186						
QY	121	KRIPNKKPGSKYTTKPKKPKPTTKKDKLDQDTKPKKEVPTTKPTKEPTINTTKTNIITT	180						
Db	187	KRIPNKKPGSKYTTKPKKPKPTTKKDKLDQDTKPKKEVPTTKPTKEPTINTTKTNIITT	246						
QY	181	LLNNNTTGNPKLISQMEFTHSSSGSNI.SPQVSTTSHSPQSSPNNTRQ	232						
Db	247	LLNSNTTGNPKLISQMEFTHSSSGSNI.SPQVSTTSHSPQSSPNNTRQ	298						

RESULT 7	
AAU74676	
ID	AAU74676 standard. Protein. 298 AA.
XX	
AC	AAU74676;
XX	
DT	09-APR-2002 (first entry)
DE	Respiratory syncytial virus G protein.
XX	
KW	RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;
KW	vaccine; gene therapy; paramyxovirus; sendai virus; PMV;
KW	antiviral chemotherapeutic compound; humoral response;
KW	cellular immune response; hPIV; paediatric respiratory disease;
KW	globin gene transfer; sickle cell disease; beta-thalassaemia;
KW	human immunodeficiency virus infection; HIV.
XX	
OS	Human respiratory syncytial virus.
PN	W0200192548-A2.
XX	
XX	06-DEC-2001.
PD	
XX	
PF	22-MAY-2001; 2001WO-US16610.
XX	
PR	01-JUN-2000; 2000US-208701P.
XX	
XX	
PA	(SJDJ-) ST JUDE CHILDREN'S RES HOSPITAL.
XX	
PI	Portner A, Takimoto T;
XX	
DR	WPI; 2002-130534/17.
DR	N-PSDB; MMS21045.
XX	
PT	Recombinant Sendai virus useful in vaccines to protect infection by
PT	paramyxoviruses, comprises exogenous nucleic acid encoding
PT	paramyxovirus protein or its antigenic fragment -
XX	
PS	Disclosure; Page 48; 57pp; English.
XX	
CC	The invention relates to a recombinant Sendai virus comprising an
CC	exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its
CC	antigenic fragment. The virus may be administered in combination
CC	with an antiviral chemotherapeutic compound. Two or more viruses
CC	expressing different PMV proteins may be co-administered. Compositions
CC	comprising the virus are useful for eliciting a humoral and/or
CC	cellular immune response to a PMV in a mammal, particularly a human.
CC	Further a recombinant Sendai virus comprising an exogenous nucleic acid
CC	encoding a second PMV protein is also administered and priming and/or
CC	boosting humoral or cellular immune response comprises administering
CC	one or more of a recombinant or isolated PMV protein or its antigenic
CC	fragment, a DNA vaccine encoding the same, and a non-Sendai viral
CC	vector encoding a PMV protein. The recombinant virus is useful as an
CC	effective vaccine against hPIV or RSV (the major causes of paediatric

CC respiratory disease) and also to express any gene of  
 CC interest in target cells, providing a positive medical impact on  
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 CC into stem cells effects a cure for sickle cell disease or beta-  
 CC thalassemia. The recombinant virus may also prove effective in  
 CC conferring immunity to human immunodeficiency virus (HIV) infection.  
 CC The Sendai virus replicates at level that is high enough to  
 CC induce sufficient immunity, but does not cause any harm to human  
 CC recipient. The present sequence represents a respiratory syncytial  
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 CC suitable for expression by the recombinant virus of the invention.

XX Sequence 298 AA:

Query Match 91.1%; Score 1120; DB 23; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.6e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSETTSQTTILASTPGVKSNL 60  
 Db 67 HKVPTTAIIODATSOIKNTTPYLTQNPOLGISPSNPSETTSQTTILASTPGVKS 126  
 QY 61 OPTVTKNTTTOTOPSKPTTKORONKPPNKPNNDPFEVFNVPSCISNNPTCAIC 120  
 Db 127 QSTVTKNTTTOTOPSKPTTKORONKPPNKPNNDPFEVFNVPSCISNNPTCAIC 186  
 QY 121 KRIPNKKPGKTTTKPKPKFTTKKDLKPQTTPKREVPPTKPTPEPTINTKTNTTT 180  
 Db 187 KRIPNKKPGKTTTKPKPKFTTKKDLKPQTTPKREVPPTKPTPEPTINTKTNTTT 246  
 QY 181 LLTNTTGNKRLTSQMETFHSSTSEGNLSPSOVSTTSEHPSSPNTTRQ 232  
 Db 247 LLTNTTGNPELTSQMETFHSSTSEGNLSPSOVSTTSEHPSSPNTTRQ 298

#### RESULT 8

AAP90441  
 ID AAP90441 standard; protein; 681 AA.

XX AAP90441:  
 AC  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Chimeric human respiratory syncytial virus glycoproteins F and G.  
 XX  
 KM Chimeric polypeptide; human respiratory syncytial virus;  
 KM protein F; protein G; vaccine.  
 XX  
 OS Human respiratory syncytial virus.  
 XX  
 PN WO8905823-A.  
 PD 29-JUN-1989.  
 XX  
 PF 31-OCT-1988; 88WO-US03784.  
 XX  
 PR 23-DEC-1987; 87US-0137387.  
 XX  
 PA (UPJO ) UPJOHN CO.  
 XX  
 PI Mathen M.  
 XX  
 DR WPI; 1989-206593/28.  
 XX  
 PT Chimeric human respiratory syncytial virus polypeptides(s)  
 PT - contg. immunogenic fragments from HRSV glycoproteins  
 PT F and G, for vaccine prodn.  
 XX  
 PS Claim 3; page 47-48; 50pp; English.  
 XX  
 CC Chimeric polypeptide contg. a signal sequence and one or more  
 CC immunogenic fragments from both human respiratory syncytial virus  
 CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg

CC E. coli, Chinese hamster ovary cells, murine C127 cells and  
 CC S. frugiperda.

XX Sequence 681 AA:

Query Match 73.3%; Score 901; DB 10; Length 681;  
 Best Local Similarity 91.0%; Pred. No. 7.3e-63;  
 Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 30 QLGISFNSLSETTSQTTILASTTGVKSNLOPTVTKNTTTOTOPSKPTTKORONK 89  
 Db 490 QLGISPSNPSETTSQTTILASTTGVKSTOSTVTKNTTTOTOPSKPTTKORONK 549  
 QY 90 PNKPNNDPFEVFNVPSCISNNPTCAICRIPNKKPGKTTTKPKPKFTTKKDL 149  
 Db 550 PSKPNNDPFEVFNVPSCISNNPTCAICRIPNKKPGKTTTKPKPKFTTKKDL 609  
 QY 150 KPQTTKREVPPTKPTPEPTINTKTNTTTLLTNTTGNPKRLTSQMETFHSSTSEGNLS 209  
 Db 610 KPQTTKREVPPTKPTPEPTINTKTNTTTLLTNTTGNPELTSQMETFHSSTSEGNLS 669  
 QY 210 PSQVSTTSE 218  
 Db 670 PSQVNTISQ 678

#### RESULT 9

AAB67771  
 ID AAB67771 standard; protein; 452 AA.

XX AAB67771:  
 AC  
 XX  
 DT 11-JUN-2001 (first entry)  
 XX

DE Amino acid sequence of a fusion protein of P40 and RSV antigen.

XX Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;  
 KM RSV; RSV infection; lung; respiratory tract; vaccine.

XX Synthetic.

OS Klebsiella pneumoniae.  
 OS Respiratory syncytial virus.

XX WO200121203-A1.

PD 29-MAR-2001.

XX 22-SEP-2000; 2000WO-FR02626.

XX 23-SEP-1999; 99FR-0011888.

PA (FABR ) FABRE MEDICAMENT SA PIERRE.

PI Corvalan N, Goestch L;

DR WPI; 2001-257929/26.

DR N-PSDB; AAF80153.

PT Vaccine against respiratory syncytial virus, comprises enterobacterial  
 PT outer membrane protein and viral immunogen, provides protective  
 PT response throughout the respiratory tract

XX Example 2; Page 31-32; 39pp; French.

CC The present sequence represents a fusion protein comprising a Klebsiella  
 CC pneumoniae outer membrane protein A (OmpA) designated P40 and a  
 CC respiratory syncytial virus (RSV) antigen. Enterobacterium OmpA proteins,  
 CC associated with an immunogenic peptide from RSV are used to prepare  
 CC a nasal composition that induces a protective response, against RSV  
 CC infection in the upper and lower (lung) respiratory tract. OmpA  
 CC potentiates the immune response to some immunogenic peptides,  
 CC eliminating the need for adjuvants. The method is useful for producing  
 CC vaccines for prevention or treatment of RSV infections.

```

XX Sequence 452 AA:
SQ Query Match 46.6%; Score 572.5; DB 22; Length 452;
Best Local Similarity 76.1%; Pred. No. 3.5e-37;
Matches 108; Conservative 5; Mismatches 18; Indels 11; Gaps 2;

OY 25 LQDDPLGTSFSLSEITSOFTTIIASTTPGVKSNLQPTTVTKTKNTTQOPSKPTTKQ 84
DB 322 LAPDRVEIEVKGKYEVTQ-----PGCG-----DPTVTKTKNTTQOPSKPTTKQ 370
OY 85 RQKPPKPNNDHFENFVPCISNNPTCWAICRIIPNKKPGKTTTKPTKPTKPTK 144
DB 371 RQKPPKPNNDHFENFVPCISNNPTCWAICRIIPNKKPGKTTTKPTKPTKPTK 430
OY 145 TKKDLKPQTTKREVPPTKPTK 166
DB 431 TKDRHKPQTTKREVPPTKPD 452

RESULT 10
AAR95660
ID AAR95660 standard; Protein; 349 AA.
XX
AC AAR95660;
XX
DT 21-JAN-1997 (first entry)
XX
DE Streptococcal protein G fragment BB fused to RSV hapten G2A.
XX
KW Streptococcus; protein G; carrier protein; immunogenicity; increase;
KM enhance; vaccine; anti-viral; human serum albumin; binding;
KW immunostimulation; respiratory syncytial virus; fusion protein.
XX
OS Chimeric Streptococcus sp.
XX
OS Chimeric Respiratory Syncytial Virus.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /label= signal_peptide
FT Protein 24..242
FT /label= BB
FT 247..347
FT /label= G2A
FT /note= "residues 130-230 of RSV"
XX
PN WO9614416-A1.
XX
PD 17-MAY-1996.
XX
PE 07-NOV-1995; 95WO-FR01466.
XX
PR 07-NOV-1994; 94FR-0013310.
XX
PA (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;
PI Uhlen M;
XX
DR MPI: 1996-251766/25.
DR N-PSDB; AAT31647.
XX
PT Enhancing immunogenicity by coupling immunogen to serum
PT albumin-binding protein - useful for preparing improved vaccines,
PT e.g. against Respiratory Syncytial Virus
XX
PS Disclosure; Page 81-83; 102pp; French.
XX
CC The present sequence comprises a 219 amino acid fragment of
CC Streptococcal protein G, designated BB, fused to an immunogen,
CC designated G2A, derived from amino acids 130-230 of protein G from
CC Respiratory Syncytial Virus sub-group A or B. The BB fragment is
CC able to bind to human serum albumin and thereby enhance immunogenicity

```

```

CC of any antigen, hapten or immunogen that is covalently coupled to it.
CC In this specific example, the BB fragment was found to induce T helper
CC memory cells, leading to production of anti-G2A antibodies by stimulated
CC B cells.
XX
SQ Sequence 349 AA:
SQ Query Match 45.7%; Score 562; DB 17; Length 349;
Best Local Similarity 78.1%; Pred. No. 1.8e-36;
Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

OY 37 NLSEITSOFTT--ILASTT-----PGVKSNIQPTTVTKTKNTTQOPSKPTTKORONKP 89
DB 216 NGKTLGSETTEAVDAATARSFNPILENSW--TVTKTKNTTQOPSKPTTKORONKP 272
OY 90 PNKPNDHFENFVPCISNNPTCWAICRIIPNKKPGKTTTKPTKPTKPTK 149
DB 273 PNKPNDHFENFVPCISNNPTCWAICRIIPNKKPGKTTTKPTKPTKPTK 332
OY 150 KPQTTKREVPPTKPTK 166
DB 333 KPQTTKREVPPTKPD 349

RESULT 11
AAB68028
ID AAB68028 standard; Protein; 349 AA.
XX
AC AAB68028;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of fusion protein comprising 2 G protein fragments.
XX
KW Aliphatic ammonium salt; immunogen; antigen; syncytial virus infection;
KW G protein; fusion protein.
XX
OS Synthetic.
OS Streptococcus sp.
OS Respiratory syncytial virus.
XX
FH Key Location/Qualifiers
FT Protein 120..230
FT /note= "G protein fragment of VRS",
XX
PN FR2798292-A1.
XX
PD 16-MAR-2001.
XX
PE 09-SEP-1999; 99FR-0011284.
XX
PR 09-SEP-1999; 99FR-0011284.
XX
PA (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI Beck A, Klingner C, Nguyen TN;
XX
DR MPI: 2001-267782/28.
DR N-PSDB; AAF84711.
XX
PT use of quaternary aliphatic ammonium salt and immunogen or antigen to
PT combat respiratory syncytial virus infections
XX
PS Claim 16; Page 22-23; 35pp; French.
XX
CC The specification describes a method which uses quaternary aliphatic
CC ammonium salts together with an immunogen or antigen to treat syncytial
CC virus infections. The combination of the salt with the antigen or
CC immunogen improves immunogenicity and equilibrates the Th1/Th2 immune
CC response. The method is used for the treatment of respiratory syncytial
CC virus infections. The present sequence represents a fusion protein,
CC comprising an albumin binding domain of the G protein of Streptococcus
CC sp. fused to a G protein fragment of respiratory syncytial virus (VRS).

```



SQ Sequence 101 AA;

Matches	100; conservative	0; mismatches	1; indels	0; gaps
64 TVVTKNTTTTQTQPSKPTTKORONKPPNKPNNDDHFEVFNFPVCSICNNPTCAICKRI	123			

DB 1 TVKTNNTTTTOTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCAICKRI 60  
 OY 124 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 164  
 DB 61 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 101

## RESULT 14

AAR95616 standard; Protein; 101 AA.

AAR95616;

10-FEB-1997 (first entry)

RSV sub-group A wild type protein G residues 130-230.

Heterologous protein; cell surface; Staphylococcus xylosus; protein G;  
 Staphylococcus carnosus; respiratory syncytial virus; wild type; mutant;  
 expression plasmid; fusion protein; streptococcal.

Human respiratory syncytial virus.

WO9614418-A1.

17-MAY-1996.

07-NOV-1995; 95WO-FR01465.

07-NOV-1994; 94FR-0013309.

(FABR ) FABRE MEDICAMENT SA PIERRE.

Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;

WPI: 1996-251768/25.

DR N-PSDB; AAT27073.

Expression of respiratory syncytial virus protein G at cell surface  
 of bacteria non-pathogenic for mammals, useful in orally active  
 vaccines

Claim 12; Page 13-14; 38pp; French.

The invention relates to the prodn. of a heterologous protein on the  
 surface of a cell, pref. Staphylococcus xylosus or S. carnosus, contg. a  
 fragment of the respiratory syncytial virus comprising residues 130-230  
 of the RSV protein G. This sequence is though to be the wild type  
 residues 130-230 from the RSV subgp. A. The heterologous protein may  
 also contain the subgp. B sequence (AAT27074). These sequences may be  
 mutated to replace the Cys residues at pos. 173 and 186 by residues  
 unable to form a disulphide bridge esp. Ser residues (see AAT27075-6).  
 Other mutations include substitution of the Phe residues at pos. 163,  
 165, 168 and/or 170 by Ser residues. The sequences encoding the wild  
 type or mutant RSV residues 130-230 are inserted into plasmid  
 pSE/implBBBM to create plasmid pSE/G2BBBM (wild type sequence) or  
 pSE/G2subBBBM (mutant sequence). The sequence is produced from pSE/BBBM  
 as a fusion protein where the fusion is with the streptococcal protein G  
 serum albumin binding domain. These plasmids are introduced into  
 S. xylosus or S. carnosus for expression of the wild type or mutant  
 protein.

Sequence 101 AA;

Query Match 45.4%; Score 558; DB 17; Length 101;

Best Local Similarity 99.0%; Pred. No. 8.8e-37; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 1;

OY 64 TVKTNNTTTTOTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCAICKRI 123  
 DB 1 TVKTNNTTTTOTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCAICKRI 60

OY 124 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 164  
 DB 61 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 101

## RESULT 15

AAR97050 standard; Protein; 101 AA.

AAR97050;

20-JAN-1997 (first entry)

Respiratory syncytial virus antigenic fragment G2A.

Streptococcus; protein G; carrier protein; immunogenicity; increase;  
 enhance; vaccine; anti-viral; human serum albumin; binding;  
 immunostimulation; RSV; hapten; antigen.

Respiratory Syncytial Virus.

WO9614416-A1.

17-MAY-1996.

07-NOV-1995; 95WO-FR01466.

07-NOV-1994; 94FR-0013310.

(FABR ) FABRE MEDICAMENT SA PIERRE.

Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;

Uhlen M;

WPI: 1996-251766/25.

DR N-PSDB; AAT31649.

Enhancing immunogenicity by coupling immunogen to serum  
 albumin-binding protein - useful for preparing improved vaccines,  
 e.g. against Respiratory Syncytial Virus

Claim 21; Page 29-30; 102pp; French.

The present sequence is that of an immunogen (designated G2A) derived  
 from amino acids 130-230 of protein G from Respiratory Syncytial Virus  
 sub-group A or B. When fused to a 219 amino acid fragment of  
 Streptococcal protein G, designated BB, the immunogenicity of G2A  
 was enhanced. Specifically, the BB fragment was found to induce T  
 helper memory cells, leading to production of anti-G2A antibodies by  
 stimulated B cells. The BB fragment is able to bind to human serum  
 albumin and thereby enhance immunogenicity of any antigen, hapten or  
 immunogen that is covalently coupled to it.

Sequence 101 AA;

Query Match 45.4%; Score 558; DB 17; Length 101;

Best Local Similarity 99.0%; Pred. No. 8.8e-37; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 1;

OY 64 TVKTNNTTTTOTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCAICKRI 123  
 DB 1 TVKTNNTTTTOTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCAICKRI 60

OY 124 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 164  
 DB 61 PNKKPGKKTTPKPTKPKPTTKDKLPQTTKREVPPTKP 101

Search completed: May 1, 2003, 13:44:22  
 Job time : 29.3283 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:43:28 ; Search time 12.2566 Seconds  
(without alignments)  
556.934 Million cell updates/sec

Title: US-09-462-816-4

Sequence: 1 HKVLTITAIIDATSIQIKNT.....VSTTSEHPSPSPPTTRQ 232

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1229	100.0	298	2	US-08-467-963C-8
2	1229	100.0	298	2	US-08-838-189D-8
3	1229	100.0	298	3	US-08-852-344D-8
4	1229	100.0	298	3	US-08-344-639E-8
5	1229	100.0	298	4	US-08-467-969A-8
6	1229	100.0	298	4	US-08-467-961A-8
7	1229	100.0	298	4	US-08-001-554A-8
8	901	73.3	681	6	5194595-19
9	222	18.1	37	3	US-08-793-792-12
10	193	15.7	32	3	US-08-793-792-8
11	188	15.3	216	3	US-08-928-361B-8
12	188	15.3	1837	3	US-08-928-361B-5
13	187.5	15.3	1721	3	US-08-700-651-5
14	187.5	15.3	1721	3	US-08-928-361B-6
15	181.5	14.8	216	3	US-08-928-361B-27
16	174	14.2	28	3	US-08-793-792-4
17	173.5	14.1	249	3	US-08-700-651-15
18	170.5	14.1	249	3	US-08-928-361B-10
19	170.5	13.9	263	5	PCT-US91-08177-13
20	164.5	13.4	2476	2	US-08-276-967-2
21	162	13.2	941	4	US-07-757-022B-14
22	162	13.2	1022	4	US-07-757-022B-84
23	162	13.2	1038	4	US-07-757-022B-74
24	162	13.2	1049	4	US-07-757-022B-58
25	162	13.2	1140	4	US-07-757-022B-104
26	162	13.2	1270	4	US-07-757-022B-44
27	162	13.2	1311	4	US-07-757-022B-42

28	162	13.2	1313	4	US-07-757-022B-142	Sequence 142, App
29	162	13.2	1314	4	US-07-757-022B-50	Sequence 50, Appl
30	162	13.2	1320	4	US-07-757-022B-46	Sequence 46, Appl
31	162	13.2	1320	4	US-07-757-022B-60	Sequence 60, Appl
32	162	13.2	1354	4	US-07-757-022B-48	Sequence 48, Appl
33	162	13.2	1361	4	US-07-757-022B-40	Sequence 40, Appl
34	162	13.2	1363	4	US-07-757-022B-52	Sequence 52, Appl
35	162	13.2	1404	4	US-07-757-022B-2	Sequence 2, Appl1
36	162	13.2	1404	4	US-07-757-022B-62	Sequence 62, Appl
37	158	12.9	175	3	US-08-700-651-12	Sequence 12, Appl
38	158	12.9	175	3	US-08-928-361B-17	Sequence 17, Appl
39	153.5	12.5	878	4	US-09-556-706B-2	Sequence 2, Appl1
40	151	12.3	37	3	US-08-793-792-11	Sequence 11, Appl
41	150	12.2	30	4	US-08-836-504A-6	Sequence 6, Appl1
42	148	12.0	26	4	US-08-750-624-9	Sequence 9, Appl1
43	148	12.0	629	4	US-09-241-581B-6	Sequence 6, Appl1
44	148	12.0	629	4	US-08-265-428-6	Sequence 6, Appl1
45	148	12.0	629	5	PCT-US95-07721-6	Sequence 6, Appl1

# ALIGNMENTS

RESULT 1  
US-08-467-963C-8

Sequence 8, Application US/08467963C

Patent No. 5968776

GENERAL INFORMATION:

APPLICANT: KLEIN, Michel H

APPLICANT: DU, Run-Pan

TITLE OF INVENTION: EMASYSBYN, Mary E

TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A CHIMERIC PROTEIN WHICH CONFERNS PROTECTION AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY STRUCTURAL VIRUS

TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY STRUCTURAL VIRUS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,963C

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/838,189

FILING DATE: 16-Apr-1997

APPLICATION NUMBER: US 08/001,554

FILING DATE: 06-JAN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 298 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear



|||||  
Db 67 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 126  
Oy 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 120  
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 186  
Oy 121 KRIPNKKPGKTKTTTKPKPKTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 180  
Db 187 KRIPNKKPGKTKTTTKPKPKTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 246  
Oy 181 LTNNTTGNPKLTISOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTQ 232  
Db 247 LTNNTTGNPKLTISOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTQ 298

## RESULT 4

US-08-344-639E-8  
; Sequence 8, Application US/08344639E  
; Patent No. 6033668  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES  
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,639E  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-391 MIS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-344-639E-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2,4e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 60  
Db 67 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 126  
Oy 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 120  
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 186  
Oy 121 KRIPNKKPGKTKTTTKPKPKTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 180  
Db 187 KRIPNKKPGKTKTTTKPKPKTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 246  
Oy 181 LTNNTTGNPKLTISOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTQ 232  
Db 247 LTNNTTGNPKLTISOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTQ 298

## RESULT 5

US-08-467-969A-8  
; Sequence 8, Application US/08467969A  
; Patent No. 6168786  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,969A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bn  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-969A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2,4e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 60

Db 67 HKVLTFAIIQDATSQIKNTPTTYLQDPOLGISFSNLSEITSOITTLASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 180  
Db 187 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 246  
QY 181 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 232  
Db 247 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298

RESULT 6  
US-08-467-961A-8  
; Sequence 8, Application US/08467961A  
; Patent No. 6171783  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasashyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Slim & McBurney  
; STREET: 330 University Avenue, 6TH Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,961A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-961A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.4e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTFAIIQDATSQIKNTPTTYLQDPOLGISFSNLSEITSOITTLASTTPGVKSNL 60  
|||||

Db 67 HKVLTFAIIQDATSQIKNTPTTYLQDPOLGISFSNLSEITSOITTLASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 180  
Db 187 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 246  
QY 181 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 232  
Db 247 LITNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298

RESULT 7  
US-08-001-554A-8  
; Sequence 8, Application US/08001554A  
; Patent No. 6225091  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasashyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Slim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/001,554A  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
US-08-001-554A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.4e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTFAIIQDATSQIKNTPTTYLQDPOLGISFSNLSEITSOITTLASTTPGVKSNL 60  
Db 67 HKVLTFAIIQDATSQIKNTPTTYLQDPOLGISFSNLSEITSOITTLASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 180  
Db 187 KRIPNKKGGKTTTTPKPKPTFKTTKKDLKPQTTPKREVPPTTKPEEPTINTTKNTTTT 246

```
OY 181 LITNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 232
|||||
Db 247 LITNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 298

RESULT 8
5194595-19
: Patent No. 5194595
: APPLICANT: WATHEN, MICHAEL W.
: TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING
: IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY
: SYNCYTIAL VIRUS
: NUMBER OF SEQUENCES: 19
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/543,780
: FILING DATE: 31-OCT-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 137,387
: FILING DATE: 23-DEC-1987
: SEQ ID NO:19:
: LENGTH: 681
5194595-19

Query Match 73.3%; Score 901; DB 6; Length 681;
Best Local Similarity 91.0%; Pred. No. 9.5e-75;
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 30 QLGISFSNLSSETTSOTTTIILASTPTGVKSNLOPTTVTKNTTOTOPOSKRTTKORONKP 89
|||||
Db 490 QLGISFSNLSSETTSOTTTIILASTPTGVKSNLOPTTVTKNTTOTOPOSKRTTKORONKP 549

OY 90 PNKPNDHFHFVFNFPVPCISNNPTCWAICKRIIPNKKPGKKTTPKPKPTFKTKDL 149
|||||
Db 550 PNKPNDHFHFVFNFPVPCISNNPTCWAICKRIIPNKKPGKKTTPKPKPTFKTKDL 609

OY 150 KPQTTKREKVVTTKPTPEPTINTKTNTTTLLNNTTGNKLTISOMETFHSTSEGNLS 209
|||||
Db 610 KPQTTKREKVVTTKPTPEPTINTKTNTTTLLNNTTGNKLTISOMETFHSTSEGNLS 669

OY 210 PSQVSTSE 218
|||||
Db 670 PSQVSTSE 678

RESULT 9
US-08-793-792-12
: Sequence 12, Application US/08793792
: Patent No. 6077511
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Antigenic peptides derived from the
: TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis
: TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.
: NUMBER OF SEQUENCES: 13
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/793,792
: FILING DATE:
: CLASSIFICATION: 514
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 37 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
US-08-793-792-12

Query Match 18.1%; Score 222; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.9e-14;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 NKPNDHFHFVFNFPVPCISNNPTCWAICKRIIPNKK 127
|||||
Db 1 NKPNDHFHFVFNFPVPCISNNPTCWAICKRIIPNKK 37

RESULT 10
US-08-793-792-8
: Sequence 8, Application US/08793792
: Patent No. 6077511
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Antigenic peptides derived from the
: TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis
: TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.
: NUMBER OF SEQUENCES: 13
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/793,792
: FILING DATE:
: CLASSIFICATION: 514
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 32 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
US-08-793-792-8

Query Match 15.7%; Score 193; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 KPNDHFHFVFNFPVPCISNNPTCWAICKRI 123
|||||
Db 1 KPNDHFHFVFNFPVPCISNNPTCWAICKRI 32

RESULT 11
US-08-928-361B-8
: Sequence 8, Application US/08928361B
: Patent No. 6071518
: GENERAL INFORMATION:
: APPLICANT: Petersen, Carolyn
: TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS, ANALOGS AND FRAGMENTS
: TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, AND ANALOGS AND FRAGMENTS
: TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PETERS, VERNY, JONES & BIRKA
: STREET: 385 Sherman Avenue, Suite 6
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306-1840
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/928,361B
```





173 TKTNTITLLTNTTGNPKLTSQMETFHSTISSEGNLSPQVSTTSEHPSQPPSPNTT 230

```

search completed: May
Job time : 14.7566 secs

```



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:46:33 ; Search time 60.4075 Seconds  
(without alignments)  
331.391 Million cell updates/sec

Title: US-09-462-816-4  
Perfect score: 1229  
Sequence: 1 HKVTLTIAIQDQTSQIKNT.....VSTSEHPSPSSPNTTRQ 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	15.7	5179	US-10-025-380-1068	Sequence 1068, App
2	193	15.7	5179	US-09-922-217-1068	Sequence 1068, App
3	193	15.7	5179	US-09-833-263-1068	Sequence 1068, App
4	169.5	13.8	1367	US-09-801-368-108	Sequence 108, App
5	162	13.2	941	US-10-124-557-14	Sequence 14, App
6	162	13.2	1022	US-10-124-557-84	Sequence 84, App
7	162	13.2	1038	US-10-124-557-74	Sequence 74, App
8	162	13.2	1049	US-10-124-557-58	Sequence 58, App
9	162	13.2	1140	US-10-124-557-104	Sequence 104, App
10	162	13.2	1270	US-10-124-557-42	Sequence 42, App
11	162	13.2	1311	US-10-124-557-142	Sequence 142, App
12	162	13.2	1313	US-10-124-557-50	Sequence 50, App
13	162	13.2	1314	US-10-124-557-46	Sequence 46, App
14	162	13.2	1320	US-10-124-557-60	Sequence 60, App
15	162	13.2	1320	US-10-124-557-60	Sequence 60, App
16	162	13.2	1354	US-10-124-557-48	Sequence 48, App
17	162	13.2	1361	US-10-124-557-40	Sequence 40, App
18	162	13.2	1363	US-10-124-557-52	Sequence 52, App
19	162	13.2	1404	US-10-124-557-2	Sequence 2, App

20	162	13.2	1404	US-10-124-557-62	Sequence 62, App
21	161.5	13.1	288	US-09-216-393-341	Sequence 341, App
22	161.5	13.1	288	US-09-216-393-344	Sequence 344, App
23	144.5	11.8	2828	US-10-176-847-54	Sequence 54, App
24	144.5	11.8	2828	US-09-905-129-21	Sequence 21, App
25	144.5	11.8	2828	US-09-991-630-21	Sequence 21, App
26	143	11.6	995	US-09-984-130-48	Sequence 48, App
27	142.5	11.6	449	US-10-028-072-224	Sequence 224, App
28	142.5	11.6	449	US-10-121-049-224	Sequence 224, App
29	142.5	11.6	449	US-10-123-904-224	Sequence 224, App
30	142.5	11.6	449	US-10-140-470-224	Sequence 224, App
31	142.5	11.6	449	US-10-175-746-224	Sequence 224, App
32	142.5	11.6	449	US-10-176-918-224	Sequence 224, App
33	142.5	11.6	449	US-10-177-865-224	Sequence 224, App
34	142.5	11.6	449	US-10-177-865-224	Sequence 224, App
35	142.5	11.6	449	US-10-140-474-224	Sequence 224, App
36	142.5	11.6	449	US-10-142-431-224	Sequence 224, App
37	142.5	11.6	449	US-10-143-114-224	Sequence 224, App
38	142.5	11.6	449	US-10-140-002-224	Sequence 224, App
39	142.5	11.6	449	US-10-142-419-224	Sequence 224, App
40	142.5	11.6	449	US-10-123-262-224	Sequence 224, App
41	142.5	11.6	449	US-10-142-423-224	Sequence 224, App
42	142.5	11.6	449	US-10-121-050-224	Sequence 224, App
43	142.5	11.6	449	US-10-141-755-224	Sequence 224, App
44	142.5	11.6	449	US-10-143-032-224	Sequence 224, App
45	142.5	11.6	449	US-10-123-108-224	Sequence 224, App

## ALIGNMENTS

RESULT 1  
US-10-025-380-1068  
Sequence 1068, Application US/10025380  
Publication No. US20020182191A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Lodes, Michael J.  
APPLICANT: Secrist, Heather  
APPLICANT: Benson, Darin R.  
APPLICANT: Meagher, Madeleine Joy  
APPLICANT: Stolk, John A.  
APPLICANT: Wang, Tongtong  
APPLICANT: Jiang, Yugu  
APPLICANT: Smith, Carole L.  
APPLICANT: King, Gordon E.  
APPLICANT: Wang, Aijun  
APPLICANT: Clapper, Jonathan D.  
APPLICANT: Skeiky, Yasir A. W.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick Thomas S.  
TITLE OF INVENTION: COMPONENTS FOR IMMUNOTHERAPY AND DIAGNOSIS  
FILE REFERENCE: OF COLON CANCER AND METHODS FOR THEIR USE  
CURRENT APPLICATION NUMBER: US/10/025,380  
CURRENT FILING DATE: 2001-12-19  
NUMBER OF SEQ ID NOS: 1129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1068  
LENGTH: 5179  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-025-380-1068

Query Match 15.7% Score 193, DB 9, Length 5179;  
Best Local Similarity 30.6% Pred. No. 6.5e-06;  
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

Oy 5 LFTAIQDA--TSQIKNTTP-----TYIQDPQIGISFNSLSEIT--SQTTLIASTPG 55  
Db 1504 MTPPTTPASTITLTPPTTTPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563

```
QY 56 ----VKSNIQPTTVTKNTTT-----OTQPSKPTTKQKONKPKPNNDHF 98
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1564 PPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 99 FEVFNVPSCISNNPCTCAICKRIPNKKPKGKTTTKPKTKPTTKKDLKPQTKPKE 158
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1624 TPT--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 159 VPTTKPEEPTINTKTNTITLLTNNTGNPKLTSQMETFHTSSEGLSPSOVSTSE 218
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 219 HPSQSPSPNTT 230
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1732 SPPTTTPSPPTT 1743
      : 111 : 1111 : 11111 : 11111 : 11111 :

RESULT 2
US-09-922-217-1068
: Sequence 1068, Application US/09922217
: Patent No. US2002007641A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Lodes, Michael J.
: APPLICANT: Secrist, Heather
: APPLICANT: Benson, Darin R.
: APPLICANT: Meagher, Madeleine Joy
: APPLICANT: Stolk, John A.
: APPLICANT: Wang, Tongtong
: APPLICANT: Jiang, Yugu
: APPLICANT: Smith, Carole Lynn
: APPLICANT: King, Gordon E.
: APPLICANT: Wang, Aijun
: APPLICANT: Clapper, Jonathan D.
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
: TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
: FILE REFERENCE: 210121.471C13
: CURRENT APPLICATION NUMBER: US/09/922,217
: CURRENT FILING DATE: 2001-08-03
: NUMBER OF SEQ ID NOS: 1124
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1068
: LENGTH: 5179
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-922-217-1068

Query Match 15.7%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.5e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLQDPOLGISFSNLSKIT--SQTITLASTTPG 55
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1504 MTTPTTPPASTTTLPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 56 ----VKSNIQPTTVTKNTTT-----OTQPSKPTTKQKONKPKPNNDHF 98
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1564 PPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 99 FEVFNVPSCISNNPCTCAICKRIPNKKPKGKTTTKPKTKPTTKKDLKPQTKPKE 158
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1624 TPT--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 159 VPTTKPEEPTINTKTNTITLLTNNTGNPKLTSQMETFHTSSEGLSPSOVSTSE 218
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 219 HPSQSPSPNTT 230
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1732 SPPTTTPSPPTT 1743
      : 111 : 1111 : 11111 : 11111 : 11111 :

RESULT 3
```

```
US-09-833-263-1068
: Sequence 1068, Application US/09833263
: Patent No. US20020110547A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Aijun
: APPLICANT: Clapper, Jonathan D.
: APPLICANT: Stolk, John A.
: APPLICANT: Meagher, Madeleine J.
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
: FILE REFERENCE: 210121.471C12
: CURRENT APPLICATION NUMBER: US/09/833,263
: CURRENT FILING DATE: 2001-04-10
: NUMBER OF SEQ ID NOS: 1093
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1068
: LENGTH: 5179
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-833-263-1068

Query Match 15.7%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.5e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLQDPOLGISFSNLSKIT--SQTITLASTTPG 55
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1504 MTTPTTPPASTTTLPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 56 ----VKSNIQPTTVTKNTTT-----OTQPSKPTTKQKONKPKPNNDHF 98
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1564 PPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 99 FEVFNVPSCISNNPCTCAICKRIPNKKPKGKTTTKPKTKPTTKKDLKPQTKPKE 158
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1624 TPT--TPTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1675
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 159 VPTTKPEEPTINTKTNTITLLTNNTGNPKLTSQMETFHTSSEGLSPSOVSTSE 218
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1676 PPTTTPSS--PITTTSPPTTTPPTTTPPTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      : 111 : 1111 : 11111 : 11111 : 11111 :
QY 219 HPSQSPSPNTT 230
      : 111 : 1111 : 11111 : 11111 : 11111 :
Db 1732 SPPTTTPSPPTT 1743
      : 111 : 1111 : 11111 : 11111 : 11111 :

RESULT 4
US-09-801-368-108
: Sequence 108, Application US/09801368
: Patent No. US20020128250A1
: GENERAL INFORMATION:
: APPLICANT: Busby, Robert
: APPLICANT: Cali, Brian
: APPLICANT: Hecht, Peter
: APPLICANT: Holtzman, Doug
: APPLICANT: Madden, Kevin
: APPLICANT: Maxon, Mary
: APPLICANT: Milne, Todd
: APPLICANT: No. US20020128250A1man, Thea
: APPLICANT: Royer, John
: APPLICANT: Salama, Sofie
: APPLICANT: Sherman, Amir
: APPLICANT: Silva, Jeff
: APPLICANT: Summers, Eric
: TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
: FILE REFERENCE: 109272.147
: CURRENT APPLICATION NUMBER: US/09/801,368
: CURRENT FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: US 09/487,558
: PRIOR FILING DATE: 2000-01-19
: PRIOR APPLICATION NUMBER: US 60/160,587
: PRIOR FILING DATE: 1999-10-20
: NUMBER OF SEQ ID NOS: 440
```





RESULT 10-557-44  
US-10-124-557-44  
Sequence 44, Application US/10124557  
Patent No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gessner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
City: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:





```

      TITLE OF INVENTION: Megakaryocyte Stimulating Factors
      NUMBER OF SEQUENCES: 143
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Genetics Institute, Inc.
      STREET: 87 Cambridgepark Drive
      CITY: Cambridge
      STATE: Massachusetts
      COUNTRY: U.S.A.
      ZIP: 02140
      COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/10/124,557
      FILING DATE: 16-Apr-2002
      CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/643,502
      FILING DATE: 18-JAN-1991
      APPLICATION NUMBER: US 07/546,114
      FILING DATE: 29-JUN-1990
      APPLICATION NUMBER: US 07/457,196
      FILING DATE: 29-DEC-1989
      APPLICATION NUMBER: US 07/390,901
      FILING DATE: 08-AUG-1989
      ATTORNEY/AGENT INFORMATION:
      NAME: Cserr, Luann
      REGISTRATION NUMBER: 31,822
      REFERENCE/DOCKET NUMBER: GI 5190
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)876-1170
      TELEFAX: (617)876-5851
      INFORMATION FOR SEQ ID NO: 142:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1313 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-124-557-142

Query Match      13.2%; Score 162; DB 12; Length 1313;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIIQDATSQIKNTPTTYLTDDPOLGISFSNLSEITSOITTIILASTPGVKSNIQPTTV 65
DB 233 TSKVLAKPTPKAEITTK-----GPAITTPKEPTPTTKEPASTTP---KEPTPTTI 280
QY 66 K-----TKNTTTOPTOPSKPTTKQORONKPNPNDFHEVFNFPSCISN 112
DB 281 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTKEP-----APTTPKEP 329
QY 113 NPTCAIICKRIPN-----KKPGKKTTPK-----TKKPTFKTTK 146
DB 330 APTTTKSAPTTPKEPAPTTKKPAPTTPKEPAPTTKKEPAPTTKKEPAPTTK 389
QY 147 KDL-----KPTTKKEVPPTTK-----TEEPTINTTKNTNTTLLTNNTTGNPKLTS 194
DB 390 EPAPTAPKPAPTTPKKEPAPTTKKEPAPTTKKEPSPTTPKKEPAPTTKSAPTTKKEPAPT 449
QY 195 QMETFHSISSEGNLSPSOVSTTSEHPSOPSSPNTT 230
DB 450 TTKSAPTTPK-----PSPPTTKKEPAPTTKKEPAPTT 481

```

## RESULT 13

```

US-10-124-557-50
Sequence 50, Application US/10124557
Patent No. US20020137894A1
GENERAL INFORMATION:

```

```

      APPLICANT: Turner, Katherine
      Clark, Stephen C.
      Jacobs, Kenneth
      Hewick, Rodney M.
      Gesner, Thomas G.
      TITLE OF INVENTION: Megakaryocyte Stimulating Factors
      NUMBER OF SEQUENCES: 143
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Genetics Institute, Inc.
      STREET: 87 Cambridgepark Drive
      CITY: Cambridge
      STATE: Massachusetts
      COUNTRY: U.S.A.
      ZIP: 02140
      COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/10/124,557
      FILING DATE: 16-Apr-2002
      CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/643,502
      FILING DATE: 18-JAN-1991
      APPLICATION NUMBER: US 07/546,114
      FILING DATE: 29-JUN-1990
      APPLICATION NUMBER: US 07/457,196
      FILING DATE: 29-DEC-1989
      APPLICATION NUMBER: US 07/390,901
      FILING DATE: 08-AUG-1989
      ATTORNEY/AGENT INFORMATION:
      NAME: Cserr, Luann
      REGISTRATION NUMBER: 31,822
      REFERENCE/DOCKET NUMBER: GI 5190
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)876-1170
      TELEFAX: (617)876-5851
      INFORMATION FOR SEQ ID NO: 50:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1314 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-124-557-50

Query Match      13.2%; Score 162; DB 12; Length 1314;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIIQDATSQIKNTPTTYLTDDPOLGISFSNLSEITSOITTIILASTPGVKSNIQPTTV 65
DB 234 TSKVLAPTPKAEITTK-----GPAITTPKEPTPTTKEPASTTP---KEPTPTTI 281
QY 66 K-----TKNTTTOPTOPSKPTTKQORONKPNPNDFHEVFNFPSCISN 112
DB 282 KSAPTTPKEPAPTTKKAPPTTKKEPAPTTKKEPAPTTKEP-----APTTPKEP 330
QY 113 NPTCAIICKRIPN-----KKPGKKTTPK-----TKKPTFKTTK 146
DB 331 APTTTKSAPTTPKEPAPTTKKPAPTTPKEPAPTTKKEPAPTTKKEPAPTTK 390
QY 147 KDL-----KPTTKKKEVPPTTK-----TEEPTINTTKNTNTTLLTNNTTGNPKLTS 194
DB 391 EPAPTAPKPAPTTPKKEPAPTTKKEPAPTTKKEPSPTTPKKEPAPTTKSAPTTKKEPAPT 450
QY 195 QMETFHSISSEGNLSPSOVSTTSEHPSOPSSPNTT 230
DB 451 TTKSAPTTPK-----PSPPTTKKEPAPTTKKEPAPTT 482

```

RESULT 14  
US-10-124-557-46  
Sequence 46, Application US/10124557  
Patent No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/124,557  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
US-10-124-557-46  
Query Match 13.2%; Score 162; DB 12; Length 1320;  
Best Local Similarity 24.6%; Pred. No. 0.0003;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIIQDATSQIKNTTPYTLTODPOLGISFSNLSITQSOTTTILASTTPGVKSNLOPTTV 65  
DB 240 TSKVLAKPPTKPAETTK-----GPAITPKKEPTPTTPKEPASTTP--KEPTPTTI 287  
QY 66 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDPFHFVENFVPCISCSN 112  
DB 288 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTKPEP-----APTTTKEP 336  
QY 113 NPTCAICKRIPN-----KKPGKTTTKP-----TKKPTTKTK 146  
DB 337 APPTTKSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTKKEPAPTTK 396  
QY 147 KDL-----KQGTTPKKEVPTTKP-----TEEPTINTTKNTITTLINNTTGNFKLTS 194  
DB 397 EPAPTPAKKPAPTTPKEPAPTTKKEPAPTTKKEPAPTTKKEPAPTTKSAPTTKKEPAPT 456

QY 195 QMETPHSTSEGNLSPSOVSTSEHPSPSPPTTV 230  
DB 457 TTKSAPTTPKE-----PSPITTKKEPAPTTKKEPAPT 488  
RESULT 15  
US-10-124-557-60  
Sequence 60, Application US/10124557  
Patent No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/124,557  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 60:  
US-10-124-557-60  
Query Match 13.2%; Score 162; DB 12; Length 1320;  
Best Local Similarity 24.6%; Pred. No. 0.0003;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIIQDATSQIKNTTPYTLTODPOLGISFSNLSITQSOTTTILASTTPGVKSNLOPTTV 65  
DB 240 TSKVLAKPPTKPAETTK-----GPAITPKKEPTPTTPKEPASTTP--KEPTPTTI 287  
QY 66 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDPFHFVENFVPCISCSN 112  
DB 288 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTKPEP-----APTTTKEP 336  
QY 113 NPTCAICKRIPN-----KKPGKTTTKP-----TKKPTTKTK 146  
DB 337 APPTTKSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTKKEPAPTTK 396

```
OY 147 KDL-----KQYTKPKREVTTKP-----TEEPTINTKTNTTLLTNNNTGNPKLTS 194
      :      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 397 EPAPTAPEKAPATTTPKEPAPTTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTKEPAPT 456
      :      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 195 QMETFHSTSSBGNLSPSOVSTTSEHPSQSPSPNTT 230
      :      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 457 TTSAPTTPKE-----PSPTTTPKEPAPTTTPKEPAPTT 488
      :      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 1, 2003, 13:56:40  
Job time : 66.4075 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:41:33 ; Search time 15.3208 Seconds  
(Without alignments)  
1455.750 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKYLTITTAITODATSOIKNT.....VSTTSEHPSPSSPNTTRQ 232

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1216	98.9	298	1 MGNZRL	major surface glycoprotein
2	1120	91.1	298	1 MGNZ	major surface glycoprotein
3	1052	85.6	297	2 J01208	attachment protein
4	1045	85.0	297	2 J01205	attachment protein
5	1042	84.8	298	2 J05680	G protein - Human
6	1037	84.4	297	2 J01209	attachment protein
7	1033	84.1	297	2 J01206	attachment protein
8	1019	82.9	298	2 J01207	attachment protein
9	1009	82.1	297	2 J01204	attachment protein
10	505.5	41.1	292	1 MGNZ60	major surface glycoprotein
11	497	40.4	292	1 MGNZ18	major surface glycoprotein
12	425	34.6	467	1 VHNZ	nucleocapsid protein
13	202	16.4	307	1 GSPF3	salivary glue protein
14	193	15.7	3020	2 A43932	mucin 2 precursor
15	191	15.5	662	2 A45155	mucin FIM-C.1 - Af
16	190	15.5	332	2 A53715	apomucin precursor
17	188	15.3	1832	2 T31113	mucin-like glycoprotein
18	186.5	15.2	851	2 T22696	hypothetical protein
19	185.5	15.1	379	2 S50125	salivary glue protein
20	183	14.9	217	2 S01358	salivary glue protein
21	183	14.9	1161	2 S57180	probable membrane protein
22	179	14.6	354	2 T46740	microfilament sheath
23	178	14.5	250	2 P00768	glycoprotein G - b
24	177	14.4	257	1 MGNZBR	major surface glycoprotein
25	175	14.2	770	2 T22808	hypothetical protein
26	175	14.2	825	2 T29634	hypothetical protein
27	172.5	14.0	327	2 S20074	promastigote surface
28	171.5	14.0	216	2 I51920	mucin - rhesus mac
29	170.5	13.9	263	2 J02284	glycoprotein G - b

30	170.5	13.9	279	2 S53363	mucin 5AC (clone J
31	169.5	13.8	1367	1 S44478	glycan 1.4-alpha-g
32	164.5	13.4	2476	2 T34022	zonadhesin - pig
33	163.5	13.3	294	2 A37232	mucin, tracheal (A
34	163.5	13.3	402	2 E86185	hypothetical prote
35	163	13.3	371	2 S20075	promastigote surfa
36	163	13.3	477	2 S53362	mucin 5AC (clone J
37	160.5	13.1	393	2 S63335	771-7 protein - fr
38	160	13.0	263	2 S01360	salivary glue prot
39	159.5	13.0	263	2 J02388	glycoprotein G - o
40	158	12.9	1777	2 T34369	hypothetical prote
41	157.5	12.8	1630	2 A53577	ascites staloglyco
42	157	12.8	248	2 P00769	glycoprotein G - b
43	156.5	12.7	263	2 A48732	attachment glycopro
44	155	12.6	1118	2 A48292	mucin, tracheobron
45	153.5	12.5	3507	2 T34513	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

MGNZRL  
Major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: A32703; S12279

R:Johnson, P. R.; Spriggs, M. K.; Olmsted, R. A.; Collins, P. L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: A32703; MUID:87289657; PMID:2441388

A:Accession: A32703

A:Molecule type: mRNA

A:Residues: 1-298 <IOH>

A:Cross-references: GB:M17212; NID:9333940; PIDN:AAA4741.1; PID:9333941

R:Garcia-Barrero, B.; Portela, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.

EMBO J. 9, 4181-4187, 1990

A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizat

A:Reference number: S12279; MUID:91065351; PMID:2249671

A:Accession: S12279

A:Molecule type: mRNA

A:Residues: 1-298 <GAR>

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:65,103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #sta

##### Query Match

98.9%; Score 1216; DB 1; Length 298;

Best Local Similarity 99.1%; Pred. No. 6, 2e-75;

Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	HKYLTITTAITODATSOIKNTPTTYLTODPOLGTSFNSLSEITSOYTTILASTTPGVKSL	60
DB	67	HKYLTITTAITODATSOIKNTPTTYLTODPOLGTSFNSLSEITSOYTTILASTTPGVKSL	126
QY	61	QPTTVTKNTTQTOPSKPTTKORONKPPKPNFHEVEFNFCISCSNNPTCAIC	120
DB	127	QPTTVTKNTTQTOPSKPTTKORONKPPKPNFHEVEFNFCISCSNNPTCAIC	186
QY	121	KRIPNKPKGKTTTKPTKKPTKTKKDKLPQTKRKEVPTTKPEEPITNTTKINITTT	180
DB	187	KRIPNKPKGKTTTKPTKKPTKTKKDKLPQTKRKEVPTTKPEEPITNTTKINITTT	246
QY	181	LTNNNTGNPKLTSOMETHSTSEGNLSPOVSTTSEHPSPSSPNTTRQ	232
DB	247	LTNNNTGNPKLTSOMETHSTSEGNLSPOVSTTSEHPSPSSPNTTRQ	298

##### RESULT 2

MGNZ  
Major surface glycoprotein G - human respiratory syncytial virus

C:Species: human respiratory syncytial virus

C>Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 24-Sep-1999



F:67-298/Domain: extracellular #status predicted <EXC>

Query Match

Best Local Similarity 84.8%; Score 1042; DB 2; Length 298;

Matches 201; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

```
OY 1 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNL 60
    |||||||
DB 67 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNT 126
OY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 120
    |||||||
DB 127 OSTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 186
OY 121 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVPTTKPEEPTINTTKNTITT 180
    |||||||
DB 187 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVLTTPKEKPTINTTKNTIRT 246
OY 181 LITNNTGNPKLTSOMETFHSTSSGNSLSPSOVSTTSEHPSPSSPNTTQ 232
    |||||||
DB 247 LITNNTGNPEHTSQKETLHSTSEGNPSPOVYTTSEVLSQSLSPSNT 298
```

#### RESULT 6

JO1209

Attachment protein - human respiratory syncytial virus (strain RSB614)

N:Alternate names: G protein

C:Species: human respiratory syncytial virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: JO1209

R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.

J. Gen. Virol. 72, 2091-2096, 1991

A:Title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: JO1204; PMID:91374005; PMID:1895054

A:Accession: JO1209

A:Molecule type: mRNA

A:Residues: 1-297 <CAN>

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract illness in children and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 84.4%; Score 1037; DB 2; Length 297;

Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

```
OY 1 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNL 60
    |||||||
DB 67 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNT 126
OY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 120
    |||||||
DB 127 OSTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 186
OY 121 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVPTTKPEEPTINTTKNTITT 180
    |||||||
DB 187 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVLTTPKEKPTINTTKNTIRT 246
OY 181 LITNNTGNPKLTSOMETFHSTSSGNSLSPSOVSTTSEHPSPSSPNTTQ 230
    |||||||
DB 247 LITNNTGNPEHTSQKETLHSTSEGNPSPOVYTTSEVLSQSLSPSNT 296
```

#### RESULT 7

JO1206

Attachment protein - human respiratory syncytial virus (strain RSB657)

N:Alternate names: G protein

C:Species: human respiratory syncytial virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: JO1206

R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.

J. Gen. Virol. 72, 2091-2096, 1991

A:Title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: JO1204; PMID:91374005; PMID:1895054

A:Accession: JO1206

A:Molecule type: mRNA

A:Residues: 1-297 <CAN>

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract illness in children and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 84.1%; Score 1033; DB 2; Length 297;

Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

```
OY 1 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNL 60
    |||||||
DB 67 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNT 126
OY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 120
    |||||||
DB 127 OSTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 186
OY 121 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVPTTKPEEPTINTTKNTITT 180
    |||||||
DB 187 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVLTTPKEKPTINTTKNTIRT 246
OY 181 LITNNTGNPKLTSOMETFHSTSSGNSLSPSOVSTTSEHPSPSSPNTTQ 230
    |||||||
DB 247 LITNNTGNPEHTSQKETLHSTSEGNPSPOVYTTSEVLSQSLSPSNT 296
```

#### RESULT 8

JO1207

Attachment protein - human respiratory syncytial virus (strain XSB6190)

N:Alternate names: G protein

C:Species: human respiratory syncytial virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: JO1207

R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.

J. Gen. Virol. 72, 2091-2096, 1991

A:Title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: JO1204; PMID:91374005; PMID:1895054

A:Accession: JO1207

A:Molecule type: mRNA

A:Residues: 1-298 <CAN>

A:Note: the authors translated the codon ACC for residue 4 as Asn and ACC for residue 11 as Lys.

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract illness in children and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,237,250,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 82.9%; Score 1019; DB 2; Length 298;

Matches 197; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

```
OY 1 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNL 60
    |||||||
DB 67 HKVLTALIQDTSQIKNTPTTYLTQDPOLGIFSNSLSEITTSQTTTILASTPGVKSNT 126
OY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 120
    |||||||
DB 127 OSTVTKTKNTTTTQTOPSKPTTKORONKPPNNDHFHFVFNPCISCSNPTCMAIC 186
OY 121 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVPTTKPEEPTINTTKNTITT 180
    |||||||
DB 187 KRIPNKKPGKKTTPPKPKPTTKKDLKPQTTKPEVLTTPKEKPTINTTKNTIRT 246
OY 181 LITNNTGNPKLTSOMETFHSTSSGNSLSPSOVSTTSEHPSPSSPNTTQ 232
    |||||||
DB 247 LITNNTGNPEHTSQKETLHSTSEGNPSPOVYTTSEVLSQSLSPSNT 298
```

## RESULT 9

J01204 attachment protein - human respiratory syncytial virus (strain RSB642)

M:Alternate names: G protein

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: J01204

R:Canv, P.A.; Matthews, D.A.; Pringle, C.R.

J Gen. Virol. 72, 2091-2096, 1991

A:title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: J01204; MUID:91374005; PMID:1895054

A:Accession: J01204

A:Molecule type: mRNA

A:Residues: 1-297 &lt;CAN&gt;

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill

children and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 1009; DB 2; Length 297;

Best Local Similarity 84.0%; Pred. No. 5.1e-61;

Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

QY 1 HKVLTITAIQDATSQIKNTPTTYLTQDPQGISFSNLSSETTSQTTTILASTTPGVKSNL 60

DB 67 HRTVSTTTIIQDATNQIKNTPTTYLTQDPQGISFSNLSSETTSQTTTILASTTPGVKSNL 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKPNNDHFEEVFNVPSCISNNPTCMAIC 120

DB 127 QSTTVRKNTTTTQADPNKSTKQKRONKPPKPNNDHFEEVFNVPSCISNNPTCMAIC 186

QY 121 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 180

DB 187 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 246

QY 181 LTLNNTGNPKLTQSOMETFSSTSEGNLSPSOVSTSEHPSQSPSPNTR 231

DB 247 PLTNSNTRNDELTSOMETFSSTSEGNLSPSOVSTSEHPSQSPSPNTR 297

## RESULT 10

MGN260

Major surface glycoprotein G - human respiratory syncytial virus (strain 8/60)

N:Alternate names: attachment glycoprotein G

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C:Accession: A37077

R:Sullender, W.M.; Anderson, K.; Wertz, G.W.

Virology 178, 195-203, 1990

A:title: The respiratory syncytial virus subgroup B attachment glycoprotein: analysis of

gous subgroup virus challenge.

A:Reference number: A37077; MUID:90357765; PMID:1697126

A:Accession: A37077

A:Molecule type: mRNA

A:Residues: 1-292 &lt;SUL&gt;

A:Cross-references: EMBL:M55633; NID:g333944; PIDN:AAA47413.1; PID:g333945

C:Genetics:

A:Gene: G

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:45-63/Domain: transmembrane #status predicted &lt;TMN&gt;

F:81,86,100,230,290/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 41.1%; Score 505.5; DB 1; Length 292;

Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;

QY 1 HKVLTITAIQDATSQIKNTPTTYLTQDPQGISFSNLSSETTSQTTTILASTTPGVKSNL 60

DB 67 HKVLTITVYQITKNHGNKISTYLTQVPPERVNSSKQPTTSPHTNSKATISPNKSET 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKPNNDHFEEVFNVPSCISNNPTCMAIC 120

DB 127 HHTTAQTKGRITTSQTKPKSTKRSKNPKKPDVDFEEVFNVPSCIGNOLCKSIC 186

QY 121 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 179

DB 187 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 246

QY 180 TLLNNTGNPKLTQSOMETFSSTSEGNLSPSOVSTSEHPSQSPSPNTR 226

DB 247 SOSTVLDITTPKTYLTQDQSLHSTSENPSTQIPTASE-PS-1SNP 251

## RESULT 11

MGN218

Major surface glycoprotein G - human respiratory syncytial virus (strain 18537)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: B32703

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

F:41-63/Domain: transmembrane #status predicted &lt;TMN&gt;

F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 497; DB 1; Length 292;

Best Local Similarity 46.8%; Pred. No. 1.3e-26;

Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;

QY 1 HKVLTITAIQDATSQIKNTPTTYLTQDPQGISFSNLSSETTSQTTTILASTTPGVKSNL 60

DB 67 HKVLTITVYQITKNHGNKISTYLTQVPPERVNSSKQPTTSPHTNSKATISPNKSET 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKPNNDHFEEVFNVPSCISNNPTCMAIC 120

DB 127 HHTTAQTKGRITTSQTKPKSTKRSKNPKKPDVDFEEVFNVPSCIGNOLCKSIC 186

QY 121 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 179

DB 187 KRIPNKKPKGKRTTKPTKPKTKKDLKPKTKPREVPTTPTEPPTINTKTNTTT 246

QY 180 TLLNNTGNPKLTQSOMETFSSTSEGNLSPSOVSTSEHPS 221

DB 247 SOSTVLDITTPKTYLTQDQSLHSTSENPSTQIPTASE-PS 287

## RESULT 12

VHNZ

nucleocapsid protein (version 2) - human respiratory syncytial virus

C:Species: human respiratory syncytial virus

C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999

C:Accession: A04026

R:Elango, N.; Venkatesan, S.

Nucleic Acids Res. 11, 5941-5951, 1983

A:title: Amino acid sequence of respiratory syncytial virus capsid protein.

A:Reference number: A04026; MUID:83299261; PMID:6310521

A:Accession: A04026

A:Molecule type: mRNA

A:Residues: 1-467 &lt;ELA&gt;

A:Cross-references: GB:X00001; NID:g61215; PIDN:CAA24906.1; PID:g61216

C:Genetics:

A:Gene: N

C:Superfamily: respiratory syncytial virus nucleocapsid protein

C:Keywords: nucleocapsid

Query Match

Best Local Similarity 34.6%; Score 425; DB 1; Length 467;

Matches 34; Conservative 11; Mismatches 12; Indels 1; Gaps 1;





Db 1564 PPPTTTTTPPTTSPPTTTTSPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623  
QY 99 FEVFNFPVPCISICNNPTCWAICRIIPNKKPGKTTTKPKPKTKKDKLPQTTPKE 158  
Db 1624 TPI--TPTSTTTLPPT-----TTPSPPTTTTTPPTTTPSPPTTTPPTTTPPTT 1675  
QY 159 VPTTKPTEPPTINTTKITTTTLLNNTGNPKLTSQMEFHTSSSEGNLSPSQVSTTSE 218  
Db 1676 PPTTPSS--PITTPSPPTTMTTPSPPTTP--SSPITTTTPSSSTTTPSPPTTMTTP 1731  
QY 219 HPSQSPSPNTT 230  
Db 1732 SPPTTSPPTT 1743

RESULT 15  
A45155

mucin FIM-C.1 - African clawed frog (fragment)  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000  
C:Accession: A45155  
R:Hauser, F.; Hoffmann, W.  
J. Biol. Chem. 267, 24620-24624, 1992  
A:Title: P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1)  
A:Reference number: A45155; MUID:93077556; PMID:1447205  
A:Accession: A45155  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-662 <HAU>  
A:Cross-References: GB:L02115; NID:9214147; PIDN:AAA74725.1; PID:9951460  
C:Superfamily: trefoil homology  
F:162-202/Domain: trefoil homology <TRP1>  
F:307-347/Domain: trefoil homology <TRP2>  
F:354-394/Domain: trefoil homology <TRP3>  
F:526-566/Domain: trefoil homology <TRP4>  
F:573-613/Domain: trefoil homology <TRP5>  
F:621-661/Domain: trefoil homology <TRP6>

## Query Match

15.5%; Score 191; DB 2; Length 662;

Best Local Similarity 29.0%; Pred. No. 1.2e-05;  
Matches 80; Conservative 19; Mismatches 111; Indels 66; Gaps 12;

QY 2 KYLTATAIQDATSOIKNTT-PTYLQDPQIGISFNSLSEITSOITTIASITPGVKSNL 60  
Db 240 KATPTTTTAKAPPTTTTAKATTTTP-----TTTTTAKATTTPTTTT 287  
QY 61 QPTTVTKNTTTTQT-----QPSKP-----TTKORNK-----PNK----- 92  
Db 288 TPPTTTTAKATTTTTSGECKMPSKREDGYSGITESOCRKGGCCPDSSIPQTKWCFTT 347  
QY 93 --PNDHFHEVFNFPVCSI-----CSNNPTC-----WAICKRIIPNKKPGKTTT 134  
Db 348 LSGVAACCKEYPSRGVCGFRGITADCGRKNCCFSSISGTKW--CFYSTQVAATKT 405  
QY 135 KPTKRPFTKTKDKLPQTTKREVPPTKPTBEPITNTTKINTTTLLNNTGNPKLTS 194  
Db 406 TPPTTTT-PTTTTAKATTTTPTTTTPTTTTPTTTTPTTTTAKATTTPTTTPTTTTAKAT 464  
QY 195 QMETFHTSSSEGNLSPSQVSTTSEHPSQSPNTT 230  
Db 465 TPPT--TTTTPTTTTAKATTT--PTTTTPTTTT 496

Search completed: May 1, 2003, 13:47:17  
Job time : 16.3208 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 12:35:37 ; Search time 10.0679 Seconds

(without alignments)  
955.759 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVLTALITADATSOIKNT.....VSTSEHPSPSPNTTRQ 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1216	98.9	298	1	VGIG_HRSVL
2	1120	91.1	298	1	VGIG_HRSVA
3	1052	85.6	297	1	VGIG_HRSV6
4	1045	85.0	297	1	VGIG_HRSV3
5	1037	84.4	297	1	VGIG_HRSV7
6	1033	84.1	297	1	VGIG_HRSV4
7	1019	82.9	298	1	VGIG_HRSV5
8	1009	82.1	297	1	VGIG_HRSV2
9	505.5	41.1	292	1	VGIG_HRSV8
10	497	40.4	292	1	VGIG_HRSV1
11	202	16.4	307	1	SGS3_DROME
12	198.5	16.2	2700	1	ZAN_HUMAN
13	193	15.7	5179	1	MOC2_HUMAN
14	191	15.5	662	1	MOC1_XENLA
15	183	14.9	217	1	SGS3_DROSI
16	183	14.9	1161	1	DAN4_YEAST
17	180.5	14.7	263	1	VGIG_YEAST
18	179.5	14.6	263	1	VGIG_HRSV1
19	177	14.4	257	1	VGIG_HRSV4
20	176.5	14.4	263	1	VGIG_HRSV3
21	170.5	13.9	263	1	VGIG_HRSV2
22	169.5	13.8	1367	1	AMYH_YEAST
23	169	13.8	257	1	VGIG_YEAST
24	166	13.5	2282	1	ZAN_RABIT
25	164.5	13.4	2476	1	ZAN_PIG
26	160	13.0	263	1	SGS3_DROYA
27	159.5	13.0	263	1	VGIG_DROSV
28	159	12.9	257	1	VGIG_HRSV5
29	156	12.7	257	1	VGIG_HRSV4
30	154	12.5	257	1	VGIG_HRSV3
31	150	12.2	3178	1	VS89_CAEEL
32	147.5	12.0	797	1	VGIX_HSVB
33	147.5	12.0	907	1	VGPI3_EBV

34	147	12.0	860	1	CH12_COICM
35	145.5	11.8	5376	1	ZAN_MOUSE
36	144.5	11.8	1260	1	ALSI_CANAL
37	143	11.6	886	1	VGPI3_EBV
38	141	11.5	1251	1	YOUN_CAEEL
39	139.5	11.4	400	1	MOC1_XENLA
40	139	11.3	767	1	AMYH_SACDI
41	138	11.2	457	1	P60_LISIN
42	137.5	11.2	681	1	VGPI3_EBV
43	135.5	11.0	768	1	AMYI_SACDI
44	135	11.0	605	1	YHCB_YEAST
45	134.5	10.9	596	1	YMAC_SCHPO

## ALIGNMENTS

```

RESULT 1
VGIG_HRSVL          STANDARD; PRT; 298 AA.
AC P20895;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup A / strain Long).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11260;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87289657; PubMed=2441388;
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;
RT "The G glycoprotein of human respiratory syncytial viruses of
RT subgroups A and B: extensive sequence divergence between
RT antigenically related proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to: license@sib-sib.ch).
CC -----
DR EMBL: M17212; AAA47411.1; -.
DR PIR: A32703; MGNZRL.
DR InterPro: IPR000925; Glycoprot G.
DR Pfam: PF00802; Glycoprotein G: 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37
FT TRANSMEM 38 66
FT DOMAIN 67 298
FT CARBOHYD 103 103 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 298 AA; 32781 MW; B79FEFA4BA73B0E CRC64;

Query Match 98.9%; Score 1216; DB 1; Length 298;
Best Local Similarity 99.1%; Pred. No. 1.3e-75;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 HKVLTAAIIOADATSOIKNTPTPTLYTODPOLGISFNSLSEITTSQTTTLLASTTPGVKSNL 60
CC |-----|
CC |-----|
CC |-----|
Db 67 HKVLTAAIIOADATSOIKNTPTPTLYTODPOLGISFNSLSEITTSQTTTLLASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 120
CC |-----|
CC |-----|
Db 127 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 186
QY 121 KRIPIKKKPKTKTTKPTKPKPTKTKDKLPQTTPKPEVPTTKPTTEPTINTKTNTTTT 180
CC |-----|
CC |-----|
Db 187 KRIPIKKKPKTKTTKPTKPKPTKTKDKLPQTTPKPEVPTTKPTTEPTINTKTNTTTT 246
QY 181 LLTNNNTGNPKLTSSOMETFHSTSSSEGNLSPSOVSTTSHPSQSSPPTTRQ 232
Db 247 LLTNNNTGNPKLTSSOMETFHSTSSSEGNLSPSOVSTTSHPSQSSPPTTRQ 298

RESULT 2
VGLG_HRSVA STANDARD; PRT; 298 AA.
ID VGLG_HRSVA
AC P03423;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain A2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85216636; PubMed=3858865;
RA Wertz G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;
RT "Nucleotide sequence of the G protein gene of human respiratory
RT syncytial virus reveals an unusual type of viral membrane protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067198; PubMed=4069997;
RA Satake M., Colligan J.E., Eliango N., Norrby E., Venkatesan S.;
RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel
RT structure.";
RL Nucleic Acids Res. 13:7795-7812(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266253; PubMed=7747420;
RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;
RT "A cold-passaged, attenuated strain of human respiratory syncytial
RT virus contains mutations in the F and L genes.";
RL Virology 208:478-484(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97187925; PubMed=9035372;
RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,
RA Murphy B.R.;
RT "Acquisition of the ts phenotype by a chemically mutagenized cold-
RT passaged human respiratory syncytial virus vaccine candidate results
RT from the acquisition of a single mutation in the polymerase (L)
RT gene.";
RL Virus Genes 13:269-273(1996).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: M11486; AAB59857.1; -
DR EMBL: X03149; CAA26928.1; -
DR EMBL: U0362; AAB86653.1; -
DR EMBL: U0363; AAB86675.1; -
DR EMBL: U0364; AAC55969.1; -
DR PIR: A04039; MGNI.
DR InterPro: IPR000925; Glycoprot-G.
DR Pfam: PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66
FT DOMAIN 67 298
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;

Query Match 91.1%; Score 1120; DB 1; Length 298;
Best Local Similarity 92.7%; Pred. No. 4e-69; Mismatches 0; Gaps 0;
Matches 215; Conservative 5; Indels 12;

QY 1 HKVLTAAIIOADATSOIKNTPTPTLYTODPOLGISFNSLSEITTSQTTTLLASTTPGVKSNL 60
Db 67 HKVLTAAIIOADATSOIKNTPTPTLYTODPOLGISFNSLSEITTSQTTTLLASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 186
QY 121 KRIPIKKKPKTKTTKPTKPKPTKTKDKLPQTTPKPEVPTTKPTTEPTINTKTNTTTT 180
Db 187 KRIPIKKKPKTKTTKPTKPKPTKTKDKLPQTTPKPEVPTTKPTTEPTINTKTNTTTT 246
QY 181 LLTNNNTGNPKLTSSOMETFHSTSSSEGNLSPSOVSTTSHPSQSSPPTTRQ 232
Db 247 LLTNNNTGNPKLTSSOMETFHSTSSSEGNLSPSOVSTTSHPSQSSPPTTRQ 298

RESULT 3
VGLG_HRSV6 STANDARD; PRT; 297 AA.
ID VGLG_HRSV6
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsh6256).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC PIR: J01208; J01208.
CC InterPro: IPR000925; Glycoprot-G.
CC Pfam: PF00802; Glycoprotein_G; 1.

```

KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 POTENTIAL.  
FT TRANSMEM 38 66 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 67 297 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA: 32708 MM: 6781756C38B64A80 CRC64;  
  
Query Match 85.6%; Score 1052; DB 1; Length 297;  
Best Local Similarity 87.4%; Pred. No. 1.5e-64;  
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
  
QY 1 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTLLASTTPGVKSNL 60  
DB 67 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTTPPAPTPSAESTP 126  
QY 61 QPTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKPKPTTKKDLKPQTTPKREVALTTKPTKEPTINTKTNIRTT 246  
QY 181 LTTNNTGNPKLTLSQMETFHSTSSSEGNLSPSOVSTTSEHPSPSPPTT 230  
DB 247 LTTNNTGNPKLTLSQMETFHSTSSSEGNLSPSOVSTTSEHPSPSPPTT 296  
  
RESULT 4  
VGIG\_HRSV3  
ID VGIG\_HRSV3 STANDARD: PRT: 297 AA.  
AC P27022;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb1734).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11253;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT \*Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.\*;  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01205; J01205.  
DR InterPro: IPR000925; Glycoprot-G.  
DR Pfam: PF00802; Glycoprotein\_G.1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA: 32525 MM: 48448F9D091E1802 CRC64;  
  
Query Match 85.0%; Score 1045; DB 1; Length 297;  
Best Local Similarity 86.6%; Pred. No. 4.5e-64;

Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;  
  
QY 1 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTLLASTTPGVKSNL 60  
DB 67 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTTPPAPTPSAESTP 126  
QY 61 QPTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKPKPTTKKDLKPQTTPKREVALTTKPTKEPTINTKTNIRTT 246  
QY 181 LTTNNTGNPKLTLSQMETFHSTSSSEGNLSPSOVSTTSEHPSPSPPTT 231  
DB 247 LTTNNTGNPKLTLSQMETFHSTSSSEGNLSPSOVSTTSEHPSPSPPTT 297  
  
RESULT 5  
VGIG\_HRSV7  
ID VGIG\_HRSV7 STANDARD: PRT: 297 AA.  
AC P27026;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb6614).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11257;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT \*Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.\*;  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01209; J01209.  
DR InterPro: IPR000925; Glycoprot-G.  
DR Pfam: PF00802; Glycoprotein\_G.1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA: 32670 MM: 58B384028E437ACD CRC64;  
  
Query Match 84.4%; Score 1037; DB 1; Length 297;  
Best Local Similarity 86.5%; Pred. No. 1.6e-63;  
Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;  
  
QY 1 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTLLASTTPGVKSNL 60  
DB 67 HKVLTALTIQDATSOIKNTTPYLYLQDPLGISFNSLSEITTSQTTTPPAPTPSAESTP 126  
QY 61 QPTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTOPSKPTTKORONKPPNNDFHEVFNFVPCISCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNITTT 180

Db 187 KRIPKKKGGKTTTKPTKKPKTKTKDKLPQTTKPEVLTTPKTEPTINTKTIRTT 246  
OY 181 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 230  
Db 247 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 296

RESULT 6  
ID VGLG\_HRSV4 STANDARD: PRT: 297 AA.  
AC P27023: (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb5857).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses."  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01206; J01206.  
DR InterPro: IPR000925; Glycoprot-G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 297 AA; 32772 MW; 10488CCAA75936BE CRC64;

Query Match 84.1%; Score 1033; DB 1; Length 297;  
Best Local Similarity 86.5%; Pred. No. 2.9e-63;  
Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

OY 1 HKVLTALTIODATSIQKNTPTVLTQDPOLGISFNSLSEITSTQTTIIASTTPGVKSNL 60  
Db 67 HKVLTALTIODATSIQKNTPTVLTQDPOLGISFNSLSEITSTQTTIIASTTPGVKSNL 126  
OY 61 OPTTVKTKNTTTQTOPSKPTTKORONKPPNKNDFEVEFNVPSCISCSNNPTCWAIC 120  
Db 127 OSTTVKTKNTTTQTOPSKPTTKORONKPPNKNDFEVEFNVPSCISCSNNPTCWAIC 186  
OY 121 KRIPNKKGGKTTTKPTKKPKTKTKDKLPQTTKPEVLTTPKTEPTINTKTIRTT 180  
Db 187 KRIPNKKGGKTTTKPTKKPKTKTKDKLPQTTKPEVLTTPKTEPTINTKTIRTT 246  
OY 181 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 230  
Db 247 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 296

RESULT 7  
VGLG\_HRSV5

ID VGLG\_HRSV5 STANDARD: PRT: 298 AA.  
AC P27024: (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb6190).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses."  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
DR PIR: J01207; J01207.  
DR InterPro: IPR000925; Glycoprot-G.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 298 AA; 32769 MW; 4D74EB54D34D7BA5 CRC64;

Query Match 82.9%; Score 1019; DB 1; Length 298;  
Best Local Similarity 84.9%; Pred. No. 2.6e-62;  
Matches 197; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

OY 1 HKVLTALTIODATSIQKNTPTVLTQDPOLGISFNSLSEITSTQTTIIASTTPGVKSNL 60  
Db 67 HKVLTALTIODATSIQKNTPTVLTQDPOLGISFNSLSEITSTQTTIIASTTPGVKSNL 126  
OY 61 OPTTVKTKNTTTQTOPSKPTTKORONKPPNKNDFEVEFNVPSCISCSNNPTCWAIC 120  
Db 127 OSTTVKTKNTTTQTOPSKPTTKORONKPPNKNDFEVEFNVPSCISCSNNPTCWAIC 186  
OY 121 KRIPNKKGGKTTTKPTKKPKTKTKDKLPQTTKPEVLTTPKTEPTINTKTIRTT 180  
Db 187 KRIPNKKGGKTTTKPTKKPKTKTKDKLPQTTKPEVLTTPKTEPTINTKTIRTT 246  
OY 181 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 232  
Db 247 LITNTTGNPKLTSMETFSHSTSGNLSPSOVSTSEHPSPSPSPPNTT 298

RESULT 8  
ID VGLG\_HRSV2 STANDARD: PRT: 297 AA.  
AC P27021: (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (strain rsb642).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11252;

[1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=91374005; PubMed=1895054;  
 RX Cane P.A., Mathews D.A., Pringle C.R.;  
 RA "Identification of variable domains of the attachment (G) protein of  
 RT subgroup A respiratory syncytial viruses.";  
 RU J. Gen. Virol. 72:2091-2096(1991).  
 CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 DR PIR: J01204; J01204.  
 DR InterPro: IPR000925; Glycoprot-G.  
 DR Pfam: PF00802; Glycoprotein.G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 297  
 FT CARBOHYD 135 135  
 FT CARBOHYD 144 144  
 FT CARBOHYD 237 237  
 SQ SEQUENCE 297 AA; 32745 MW; FC72A7F3A8B6F7C CRC64;

Query Match 82.1%; Score 1009; DB 1; Length 297;  
 Best Local Similarity 84.0%; Pred. No. 1.2e-61;  
 Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

OY 1 HKVLTLLTATIODATSOIKNTPTPYLLTODPOLGISFNSLSEITSOITTLASTTPGVKSNL 60  
 Db 67 HRYVSTTTIIODATNQIKNTPTPYLLTQNPOLGISPSNPSEITSLITLIDPTTPGVKLT 126  
 OY 61 OPTTVKTKNTTQTOPSKPTKOROKPPKPNNDHFEVFNVPVCSICSNPTCAIC 120  
 Db 127 OSTTVKTKNTTQTOPSKPTKOROKPPKPNNDHFEVFNVPVCSICSNPTCAIC 186  
 OY 121 KRIPNKKPKKRTTTPKPTKPTKDLKQTTKPKREVPPTKTEPTINTKTNTTT 180  
 Db 187 KRIPNKKPKKRTTTPKPTKPTKDLKQTTKPKREVPPTKTEPTINTKTNTTT 246  
 OY 181 LITNNTGNPKLTOSMETFHSTSEGNLSPSOVSTSEHPSQSPPTTR 231  
 Db 247 PLTSTARNDELTSOMETFHSTSEGNLSPSOVSTSEHPSQSPPTTR 297

## RESULT 9

VGIG\_HRSV8 STANDARD: PRT; 292 AA.  
 AC P23041;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 8/60).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11258;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90357765; PubMed=1697126;  
 RA Sullender W.M., Anderson K., Wertz G.W.;  
 RT "The respiratory syncytial virus subgroup B attachment glycoprotein:  
 RT analysis of sequence, expression from a recombinant vector, and  
 RT evaluation as an immunogen against homologous and heterologous  
 RT subgroup virus challenge";  
 RL Virology 178:195-203(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91374595; PubMed=1895391;  
 RA Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;

RT "Genetic diversity of the attachment protein of subgroup B  
 RT respiratory syncytial viruses.";  
 RU J. Virol. 65:5425-5434(1991).  
 CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M55633; AAA47413.1; -  
 DR EMBL: M73545; AAA7408.1; -  
 DR PIR: A37077; MGN260.  
 DR InterPro: IPR000925; Glycoprot-G.  
 DR Pfam: PF00802; Glycoprotein.G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 292  
 FT CARBOHYD 81 81  
 FT CARBOHYD 86 86  
 FT CARBOHYD 100 100  
 SQ SEQUENCE 292 AA; 32143 MW; 8EC60C85EF057BB5 CRC64;

Query Match 41.1%; Score 505.5; DB 1; Length 292;  
 Best Local Similarity 47.6%; Pred. No. 1.1e-27;  
 Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;

OY 1 HKVLTLLTATIODATSOIKNTPTPYLLTODPOLGISFNSLSEITSOITTLASTTPGVKSNL 60  
 Db 67 HRYVSTTTIIODATNQIKNTPTPYLLTQNPOLGISPSNPSEITSLITLIDPTTPGVKLT 126  
 OY 61 OPTTVKTKNTTQTOPSKPTKOROKPPKPNNDHFEVFNVPVCSICSNPTCAIC 120  
 Db 127 HHTTAQTKRITTTQTNKPSKSNPPKPKDHYHEVFNVPVCSICGNOLCKSIC 186  
 OY 121 KRIPNKKPKKRTTTPKPTKPTKDLKQTTKPKREVPPTKTEPTINTKTNTTT 179  
 Db 187 KTIPTNKKPKKRTTTPKPTKPTKDLKQTTKPKREVPPTKTEPTINTKTNTTT 246  
 OY 180 TLTNNTGNPKLTOSMETFHSTSEGNLSPSOVSTSEHPSQSPPTTR 226  
 Db 247 SSTYVITDITTPKTYTIOQSLHSTSENTPSSIOIPTASE-PS-ITSNP 291

## RESULT 10

VGIG\_HRSV1 STANDARD: PRT; 292 AA.  
 AC P20896;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 18537).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11251;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87289657; PubMed=2441388;  
 RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
 RT "The G glycoprotein of human respiratory syncytial viruses of  
 RT subgroups A and B: extensive sequence divergence between

antigenically related proteins":  
 Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  
 - FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 RESPIRATORY SYNCTIAL VIRUS G PROTEIN LACKS BOTH NEUGAMINIDASE AND  
 HEMAGGLUTININATING ACTIVITIES.  
 - SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 - PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: M17213; AAA47412.1; -  
 DR PIR: B32703; MGN218.  
 DR InterPro: IPR000925; Glycoprot.G.  
 DR Pfam: PF00802; Glycoprotein.G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 66 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 67 292  
 FT CARBOHYD 81 292 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 292 AA; 32306 MW; BC8C59F69CA7AFC2 CR64;  
 -----  
 Query Match 40.4%; Score 497; DB 1; Length 292;  
 Best local Similarity 46.8%; Pred. No. 4,1e-27;  
 Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;  
 -----  
 QY 1 HKVLTTLTAIQDARTSOIKNTTPYLYIQDPLGISFENLSEITTSQTTILASTTPGKSNL 60  
 DB 67 HKVLTTLTVYQTKNHTKNTKNTSTYLVQVPERVNSSKOPTSTPHNTSATISPNKRSST 126  
 QY 61 OPTVTKNTNTTQTOPSKFTKORONKPNPNNDFFEVNFVCSICSNPTGMAIC 120  
 DB 127 HHTTAOTKGTITSTOTNKPSTSRSKNPKKKKDYHEVFVFCISICGNOLCKSIC 186  
 QY 121 KRIPNKKPKKRTTKPKKPKFTT-KDOLKQPTKPKKPKKPKKPKKPKKPKKPKKPKK 179  
 DB 187 KTIIPNKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKKPKK 246  
 QY 180 TLITNNTGKPKKTSOMETPHSTSSSEGNLSPSGVSTSEHS 221  
 DB 247 SGSTVLDITPKYTIQOOSLHSTSTSENTSPSTOIP7ASE-PS 287  
 -----  
 RESULT 11  
 SG33\_DROME STANDARD; PRT; 307 AA.  
 AC P02840; O9VTJ2;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Salivary glue protein Sgs-3 precursor.  
 GN SG33 OR CG11720.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83294545; PubMed=6411930;  
 RA Garfunkel M.D., Pruitt R.E., Meyerowitz E.M.;  
 RT "DNA sequences, gene regulation and modular protein evolution in the  
 Drosophila 68C glue gene cluster."  
 RL J. Mol. Biol. 168:765-789(1983).

[2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Berkeley;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chert L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abdl J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandat D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
 RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,  
 RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimmel B., Kodira C.D., Kraft C., Kirovitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Munkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Sliden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 -----  
 [3]  
 RP SEQUENCE OF 1-28 FROM N.A.  
 RX MEDLINE=88332966; PubMed=3138416;  
 RA Martin G.H., Mayeda C.A., Meyerowitz E.M.;  
 RT "Evolution and expression of the Sgs-3 glue gene of Drosophila."  
 RL J. Mol. Biol. 201:273-287(1988).  
 -----  
 [4]  
 RP DEVELOPMENTAL STAGE.  
 RX MEDLINE=94038699; PubMed=8223281;  
 RA Huft F., Ruiz C., Richards G.;  
 RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
 ecdysonic responses in Drosophila."  
 RL Development 118:613-627(1993).  
 -----  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: SALIVARY GLAND SPECIFIC.  
 CC -1- DEVELOPMENTAL STAGE: IN THE SALIVARY GLANDS OF MID INSTAR LARVAE  
 CC OF DEVELOPMENT. LEVELS REMAIN CONSTANT AND ABUNDANT IN LATE LARVAE  
 CC UNTIL PUFF STAGE 10, THEN DECREASE BY STAGE 11.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: X01918; CA25394.1; -  
 DR EMBL: AE003544; AA50056.1; -  
 DR EMBL: X78392; CA55154.1; -







```
FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 4192 4192 G -> S (IN REF. 2).
SQ SEQUENCE 5179 AA; 540295 MM; 85CD57JFB9A5663 CRC64;
```

Query Match 15.78; Score 193; DB 1; Length 5179;  
Best Local Similarity 30.68; Pred. No. 2.4e-05;

Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

```
QY 5 LTTAIIODA--TSQIKNTTP-----TYLNDDPOLGISFSLSEIT--SQTITTLASTTGG 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1504 MTPPIPPASTTLPPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563
QY 56 ----VKSNIQPTTVTKTKNTTT-----OTOPSKPTTKOKRPNKPNNDPH 98
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1564 PPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
QY 99 FEVENVPVCSICNNPTCAICRIKPKGKTTTKPKTKPKTKKDLKPQTTPKE 158
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1624 TPT--TPTTSTTTLPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
QY 159 VPTTKPEPTINTTKNTITLLINNTGNPKLTQSOMFPTSTSEGNLSPSQVSTSE 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 PPTTTPSS--PTTTPSPPTTTPPTTTPPTTTP--SSPTTTPPTTTPPTTTPPTTTPPTTTP 1731
QY 219 HPSQSPSPNTT 230
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1732 SPPTTTPSPPTTTP 1743
```

```
RESULT 14
MUC1_XENLA STANDARD: PRT; 662 AA.
AC 005049;
DT 01-OCT-1994 (Rel. 30; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Integumentary mucin C.1 (FIM-C.1) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).
RC TISSUE=Skin;
RX MEDLINE=93077556; PubMed=1447205;
RA Hauser F., Hofmann W.;
RT "p-domains as shuffled cysteine-rich modules in integumentary mucin
RT C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism.";
RL J. Biol. Chem. 267:24620-24624(1992).
CC -I- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL
CC INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- ALTERNATIVE PRODUCTS: At least 7 isoforms: 1 (shown here), 2, 3,
CC 4, 5, 6 and 7; may be produced by alternative splicing.
CC -I- TISSUE SPECIFICITY: SKIN.
CC -I- PTM: EXTENSIVELY O-GLYCOSYLATED.
CC -I- SIMILARITY: CONTAINS 6 P-TYPE (TREFOL) DOMAINS.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: I02115; AAA74725.1; -.
CC PIR: A45155; A45155.
CC DR HSP: P01359; 2PSP.
CC DR InterPro: IPR000519; P_trefol.
CC DR Pfam: PF00088; trefol; 6.
CC DR SMART: SM00018; P; 6.
CC DR PROSITE: PS00025; P_TREFOL; 6.
CC DR Repeat: Amphibian skin; Glycoprotein; Alternative splicing.
CC KW NON_TER
CC FT 1 144
CC FT 81 144
CC FT 81 88
CC FT REPEAT
CC FT REPEAT 81 96
CC FT REPEAT 97 104
CC FT REPEAT 105 112
CC FT REPEAT 113 120
CC FT REPEAT 121 128
CC FT REPEAT 129 136
CC FT REPEAT 137 144
CC FT 202
CC FT 202
CC FT DOMAIN
CC FT 218
CC FT REPEAT 218 224
CC FT REPEAT 225 239
CC FT REPEAT 240 249
CC FT REPEAT 250 259
CC FT REPEAT 260 275
CC FT REPEAT 276 287
CC FT REPEAT 288 294
CC FT REPEAT 295 301
CC FT REPEAT 306 347
CC FT 353
CC FT 353
CC FT DOMAIN
CC FT 402
CC FT 402
CC FT REPEAT
CC FT REPEAT 402 411
CC FT REPEAT 412 420
CC FT REPEAT 420 431
CC FT REPEAT 432 443
CC FT REPEAT 444 453
CC FT REPEAT 454 460
CC FT REPEAT 461 472
CC FT REPEAT 473 479
CC FT REPEAT 480 491
CC FT REPEAT 492 498
CC FT REPEAT 499 515
CC FT REPEAT 516 522
CC FT REPEAT 522 525
CC FT REPEAT 525 566
CC FT DOMAIN
CC FT 572
CC FT 613
CC FT DOMAIN
CC FT 620
CC FT 661
CC FT DISULFID 172
CC FT DISULFID 188
CC FT DISULFID 187
CC FT DISULFID 182
CC FT DISULFID 172
CC FT DISULFID 307
CC FT DISULFID 333
CC FT DISULFID 332
CC FT DISULFID 344
CC FT DISULFID 354
CC FT DISULFID 380
CC FT DISULFID 364
CC FT DISULFID 374
CC FT DISULFID 379
CC FT DISULFID 391
CC FT DISULFID 526
CC FT DISULFID 536
CC FT DISULFID 551
CC FT DISULFID 546
CC FT DISULFID 573
CC FT DISULFID 583
CC FT DISULFID 593
CC FT DISULFID 610
CC FT DISULFID 621
CC FT 647
```

```
QY      84 QRC-----NKPPKNPNDDHFEVFNVPSCISNNPCLMALCKRIPKPKGOKRTTEKP 137
          |||
Db       75 SNTSTTTTTRAPPTKPT-----CKSNTSTTTTTRAPPTTKCTSTTTTTT 119
          |||
QY      138 KRPFTTKKKDLKPOTKPEVEPT--HKPEEPTINTTKTNITTT 180
          ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db       120 HKPTTHSPKTPKPTKHHTPKTKPTKHHHTPKTKPKKHHTPTTTTTT 164
```

Search completed: May 1, 2003, 13:45:04  
Job time : 12.0679 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:27:43 ; Search time 28.0151 seconds  
(without alignments)  
1706.328 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKYTLTALIQDATSIKNT.....VSTTSEHPSPSPPTTRQ 232

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL.21:\*  
2: SP\_bacteria:\*  
3: SP\_fungi:\*  
4: SP\_human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1229	100.0	298	12	Q82068	082068 human respi
2	1229	100.0	298	12	Q01929	001929 human respi
3	1093	88.9	298	12	Q09719	009719 respiratory
4	1093	88.9	298	12	Q09634	009634 human respi
5	1088	88.5	279	12	Q09678	099678 human respi
6	1072	87.2	279	12	Q09600	099600 human respi
7	1067	86.8	297	12	Q09VB5	099VB5 human respi
8	1066	86.7	297	12	Q82066	082066 human respi
9	1063	86.5	293	12	Q9YVB3	09YVB3 human respi
10	1062	86.4	298	12	Q9YVB4	09YVB4 human respi
11	1056	85.9	297	12	Q82071	082071 human respi
12	1055	85.8	278	12	Q90673	090673 human respi
13	1053	85.7	292	12	Q9YVB2	09YVB2 human respi
14	1053	85.7	297	12	Q82058	082058 human respi
15	1052	85.6	295	12	Q86359	086359 respiratory
16	1047	85.2	297	12	Q82067	082067 human respi

17	1047	85.2	297	12	Q91947	091947 human respi
18	1045	85.0	278	12	Q90671	090671 human respi
19	1045	85.0	295	12	Q86360	086360 respiratory
20	1044	84.9	279	12	Q90680	090680 human respi
21	1044	84.9	279	12	Q90689	090689 human respi
22	1044	84.9	293	12	Q9YVB1	09YVB1 human respi
23	1043	84.9	298	12	Q82065	082065 human respi
24	1041	84.7	278	12	Q90601	090601 human respi
25	1040	84.6	297	12	Q82074	082074 human respi
26	1040	84.6	297	12	Q82064	082064 human respi
27	1039	84.5	279	12	Q90652	090652 human respi
28	1039	84.5	292	12	Q9YVB0	09YVB0 human respi
29	1037	84.4	279	12	Q90659	090659 human respi
30	1037	84.4	295	12	Q86361	086361 respiratory
31	1037	84.4	297	12	Q82063	082063 human respi
32	1036	84.3	278	12	Q90674	090674 human respi
33	1036	84.3	279	12	Q90686	090686 human respi
34	1035	84.2	279	12	Q90679	090679 human respi
35	1035	84.2	299	12	Q82077	082077 human respi
36	1033	84.1	279	12	Q90657	090657 human respi
37	1033	84.1	295	12	Q86357	086357 respiratory
38	1032	84.0	279	12	Q90651	090651 human respi
39	1032	84.0	297	12	Q91946	091946 human respi
40	1032	84.0	297	12	Q9YVC8	09YVC8 human respi
41	1031	83.9	298	12	Q82060	082060 human respi
42	1029	83.7	297	12	Q82072	082072 human respi
43	1028	83.6	297	12	Q82079	082079 human respi
44	1028	83.6	297	12	Q82057	082057 human respi
45	1028	83.6	298	12	Q82062	082062 human respi

# ALIGNMENTS

## RESULT 1

ID Q82068 PRELIMINARY; PRT; 298 AA.  
AC Q82068;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Glycoprotein (Fragment).  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID-11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBGROUP A;  
RX MEDLINE-9435057; PubMed-8057427;  
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrisille S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno E.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein".  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL: Z33429; CA83872.1; -  
DR InterPro: IPR000925; Glycoprot\_G.  
DR InterPro: IPR003860; Plantine\_atlch.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
FT NON\_TER 1  
SQ SEQUENCE 298 AA; 32753 MW; 70D80307897A772B CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;  
Best Local Similarity 100.0%; Pred. No. 6.5e-85;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKYTLTALIQDATSIKNTPTLYLQDPQLGISFNSLSEITTSOTTILASTTPGVKSNL 60  
DB 67 HKYTLTALIQDATSIKNTPTLYLQDPQLGISFNSLSEITTSOTTILASTTPGVKSNL 126

	61	OPTVTKRNTKTTTQTOPSKPTTKORONKPPKNNDHFEFVNFPVCISGNNPFCMAIC	120
Dd	127	OPTVTKRNTKTTTQTOPSKPTTKORONKPPKNNDHFEFVNFPVCISGNNPFCMAIC	186
Oy	121	KRIDPKRKGGKTTTPKKPFKFTTKDKLQQTTRKPEVPRTKPEEPTINTYTNITTT	180
Dd	187	KRIENKRKGKKTTPPKPKPKFKTKDKLKQQTTRKPKPVPTTKTEEPETINTYTNTITT	246
Oy	181	LITNNTGNPKLTQSOMETFHSTSSBGNLSDPGVSTSHSPPOSSPPTTTNQ	232
Dd	247	LITNNTGNPKLTQSOMETFHSTSSBGNLSDPGVSTSHSPPOSSPPTTTNQ	298
 RESULT 2 ID ID001929 PRELIMINARY; PRT; 298 AA. 001929			
AC	001929:		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
De	Major surface glycoprotein G (Attachment glycoprotein G).		
Gn			
OS	Human respiratory syncytial virus ( subgroup A / strain Long) .		
OC	Vituses: ssRNA negative-strand viruses: Mononegavirales:		
OC	Paramyxoviridae: Pneumovirinae: Pneumovirus.		
OX	NCBI_Taxid=11260;		
RN	(1)		
Rp	SEQUENCE FROM N.A.		
Rx	MEDLINE=9106533.; PubMed=2249671;		
RA	Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;		
Rt	"Frame shift mutations as a novel mechanism for the generation of		
Rt	neutralization resistant mutants of human respiratory syncytial		
Rt	virus." ;		
RL	EMBO J. 9:4181-4187(1990).		
CC	-I- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE		
CC	RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND		
CC	HENAGGLUTININATING ACTIVITIES.		
CC	-I- CELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED		
CC	SUBCELL AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.		
CC	-I- PTM: THIS PROTEIN MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE		
CC	CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.		
CC	EMBL, X17085; CAA34937.1; ..		
DR	InterPro: IPR000925; GlycoProt-G.		
DR	InterPro: IPR003880; Pantine_attech.		
DR	Pfam: PF00802; Glycoprotein_G.1.		
DR	PROSITE: PS0012; PHOSPHOPANTHEINE; UNKNOWN_1.		
KW	Transmembrane; Glycoprotein.		
FT	DOMAIN 1 37		
FT	TRANSMEM 38 66		
FT	DOMAIN 67 298		
FT	DOMAIN 205 211		
FT	SIZE 64 75		
FT	SITE 74 75		
FT	CARBOHD 85 85		
FT	CARBOHD 103 103		
FT	CARBOHD 250 250		
FT	CARBOHD 251 251		
FT	CARBOHD 273 273		
FT	CAROHD 294 294		
FT	SEQUENCE 298 AA; 32771 MW; F8CDA213D97C2952 CRC64;		
SO			
 Query Match 100.0%; Score 1229; DB 12; Length 298; Best Local Similarity 100.0%; Pred. No. 6.5e-85; Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Oy	1	HKVTLTALIQDATSQIKNTPTLYLPDQLGISFSNLSEITSQTTTTLASTTPGVKSML	60
Dd	67	HKVTLTALIQDATSQIKNTPTLYLPDQLGISFSNLSEITSQTTTTLASTTPGVKSML	126
Oy	61	OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNNDHFEFVNFPVCISGNNPFCMAIC	120
Dd	127	OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNNDHFEFVNFPVCISGNNPFCMAIC	186

Query	121	KRIPPKKGGKTTTTPKPKPTTKTTKDKLAKOQTRKPKVPTTKPPEEPTINTKTNTTT	180			
Db	187	KRIPPKKGGKTTTTPKPKPTTKTTKDKLAKOQTRKPKVPTTKPPEEPTINTKTNTTT	246			
Qy	181	LTNNNTGNPKLTSQMETFHSTSSGKNSLSPSOVSTTSEHPSPSPNTTRQ	232			
Db	247	LTNNNTGNPKLTSQMETFHSTSSGKNSLSPSOVSTTSEHPSPSPNTTRQ	298			
RESULT 3						
ID	009719	PRELIMINARY:	PRT: 298 AA.			
AC	009719:					
DT	01-JUL-1997	(TREMBLrel. 04, Created)				
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)				
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)				
DE	Attachment glycoprotein (G).					
GN	G.					
OS	respiratory syncytial virus.					
OC	Viruses: ssRNA negative-strand viruses: Mononegavirales:					
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.					
OX	NCBI_TaxID=12814;					
RA	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-S2 TSIC;					
RX	MEDLINE=97185152; PubMed=9032893;					
RA	Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,					
RA	Longhurst S.G., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,					
RT	Pringle C.R.;					
RT	"Identification of mutations contributing to the reduced virulence of					
RL	a modified strain of respiratory syncytial virus.";					
RL	Vaccine 14:1637-1646(1996).					
RP	[2]					
RC	SEQUENCE FROM N.A.					
RC	STRAIN-S2 TSIC;					
RA	Easton A.J.;					
RL	Submitted (OCr-1995) to the EMBL/GenBank/DBJ databases.					
DR	EMBL: U39661; AAC57036.1; "					
DR	InterPro: IPR000925; Glycoprot.G.					
DR	InterPro: IPR003880; Penantre.atlch.					
DR	Pfam: PF00802; Glycoprotein_G.1					
DR	PROSITE: PS00012; PHOSPHOPANTHETHEINE, UNKNOWN_1.					
SO	SEQUENCE 298 AA: 32779 MW; 67F4A043682F4450 CRC64;					
Query Match 88.9%; Score 1093; DB 12; Length 298;						
Best Local Similarity 90.1%; Pred. No. 1.le-74;						
Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;						
Qy	1	HKVLTITAIIDATSOIKNTPTVLTODPOLGISFNSLSEITISQTTTIIASTTPGVKSNL	60			
Db	67	HKVLTITAIIDATSOIKNTPTVLTOLNPOLGISFNSLSEITISQTTTIIASTTPGVKSTL	126			
Qy	61	QPTVTKTKTTTQTOPSPKPTKONKRPKNPNDEHFEVNFVPCISCSNNPTCMAIC	120			
Db	127	QSTVTKTKTTTQTOPSPKPTKORONKRPKNPNDEHFEVNFVPCISCSNNPTCMAIC	186			
Qy	121	KRIIPNKKPKKTTTPTKPKPTTKTKDKLAPQTRKPEVPTTKPTEPTINTKTNTTT	180			
Db	187	KRIIPNKKPKKTTTPTKPKPTTKTKDKLAPQTRKPEVPTTKPTEPTINTKTNTIRT	246			
Qy	181	LTNNNTGNPKLTSQMETFHSTSSGKNSLSPSOVSTTSEHPSPSPNTTRQ	232			
Db	247	LTNNNTGNPKLTSQMETFHSTSSGKNSLSPSOVSTTSEHPSPSPNTTRQ	298			
RESULT 4						
ID	009634	PRELIMINARY:	PRT: 298 AA.			
AC	009634:					
DT	01-JUL-1997	(TREMBLrel. 04, Created)				
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)				
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)				

```

DE Attachment protein (g).
GN
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RX MEDLINE=9718152; PubMed=9032893;
RA Tolley K.P., Mariotti A.C., Simpson A., Plows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle G.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RA Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U39662; AAC57026.1; -.
DR InterPro: IPR000925; Glycoprot.G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G.1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
SQ
SEQUENCE 298 AA: 32749 MW: 60F27B29D497F31 CRC64;

Query Match 88.9%; Score 1093; DB 12; Length 298;
Best Local Similarity 90.1%; Pred. No. 11e-74;
Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 1 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNL 60
DB 67 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNTL 126
QY 61 OPTVTKKNTTTQTOTPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 120
DB 127 QSTTVTKNTTTTKTIQPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 186
QY 121 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 180
DB 187 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 246
QY 181 LITNNTGNPKLTSOMETHSTSEGNLSPSOVSTSEHPSSPNTTRQ 232
DB 247 LITNNTGNPEHTSOKETLHSTSDGNPSPSOVYTTSEYLSQPSPSNTTNG 298

RESULT 5
Q096T8 PRELIMINARY: PRT: 279 AA.
AC Q096T8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein (Fragment).
GN
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92011;
RX MEDLINE=20283719; PubMed=10823752;
RA Choi E.H., Lee H.J.;
RT "Genetic diversity and molecular epidemiology of the G protein of
RT subgroups A and B of respiratory syncytial virus isolated over 9
RT consecutive epidemics in Korea.";
RL J. Infect. Dis. 181:1547-1556(2000).
DR EMBL; AF193308; AAF23731.1; -.
DR InterPro: IPR000925; Glycoprot.G.

```

```

DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G.1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 279 AA: 30498 MW: 4E1CD2FE7E569E3A CRC64;

Query Match 88.5%; Score 1088; DB 12; Length 273;
Best Local Similarity 90.1%; Pred. No. 2.3e-74;
Matches 209; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNTL 60
DB 48 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNTL 107
QY 61 OPTVTKKNTTTQTOTPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 120
DB 108 LPTVTKKNTTTQTOTPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 167
QY 121 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 180
DB 168 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 227
QY 181 LITNNTGNPKLTSOMETHSTSEGNLSPSOVSTSEHPSSPNTTRQ 232
DB 228 LITNNTGNPEHTSOKETLHSTSDGNPSPSOVYTTSEYLSQPSPSNTTNG 279

RESULT 6
Q096U0 PRELIMINARY: PRT: 279 AA.
AC Q096U0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein (Fragment).
GN
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91242;
RX MEDLINE=20283719; PubMed=10823752;
RA Choi E.H., Lee H.J.;
RT "Genetic diversity and molecular epidemiology of the G protein of
RT subgroups A and B of respiratory syncytial virus isolated over 9
RT consecutive epidemics in Korea.";
RL J. Infect. Dis. 181:1547-1556(2000).
DR EMBL; AF193306; AAF23729.1; -.
DR InterPro: IPR000925; Glycoprot.G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G.1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 279 AA: 30544 MW: 6B270AEB1CB3533 CRC64;

Query Match 87.2%; Score 1072; DB 12; Length 273;
Best Local Similarity 90.0%; Pred. No. 3.7e-73;
Matches 207; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNTL 60
DB 48 HKVLTFTAIIDATSOIKNTPTLYLTDPOLGISFNSLSEITSOQTTLASTPGVKSNTL 107
QY 61 OPTVTKKNTTTQTOTPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 120
DB 108 LPTVTKKNTTTQTOTPSKPTTKORONKPPKPNDFHEVEFNFPVCSICSNPTGMAIC 167
QY 121 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 180
DB 168 KRIPNKKPGKKTTPKTKPKTKKDLKPDQTKPKPEVPTTKPEEPTINTTKTNITTT 227

```

Qy	181	LTNTTTCNPKLTISOMETFHSSSGCNLSPOQVSTTSEHPQSSPPMTT	230
Db	228	LTNTTTCNPKLTISOMETFHSSSGCNLSPOQVSTTSEHPQSSPPMTT	277
RESULT 7			
Q9YVB5	PRELIMINARY:	PRT:	297 AA.
AC	09YVB5:		
DT	01-MAY-1999 (TREMBLrel. 10, Created)		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
OC	Attachment glycoprotein G.		
OS	Human respiratory syncytial virus.		
OC	Viruses: ssRNA negative-strand viruses; Mononegavirales;		
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.		
OX	NCBI_TaxID=11250;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-MV2780;		
RX	MEDLINE=99022964; PubMed=9806017;		
RA	Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;		
RT	"Antigenic and genetic diversity among the attachment proteins of		
RT	group A respiratory syncytial viruses that have caused repeat		
RL	infections in children."		
DR	J. Infect. Dis. 178:925-932(1998).		
DR	EMBL; AF065405; AAD02941.1; -		
DR	InterPro: IPR000925; Glycoprot.G.		
DR	InterPro: IPR003880; Pantene_attach.		
DR	InterPro: IPR002965; P.rich.extensn.		
DR	Pfam: PF00802; Glycoprotein.G; 1.		
DR	PRINTS: PR01217; PRICHEXTENS.		
DR	PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.		
SQ	SEQUENCE 297 AA; 32572 MW; 5C10E719A05EE5C1 CRC64;		
Query Match 86.8%; Score 1067; DB 12; Length 297;			
Best local similarity 88.7%; Pred. No. 9, 4e-73;			
Matches 205; Conservative 7; Mismatches 19; Indels 0; Gaps 0.			
Qy	1	HKVLTTLAIQDATSQIKNTPTPLYTLQDPQLGISFNSLSITSQITTLTASTPGYKSNL	60
Db	67	HKVSTTTTIIQDATFNQIKNTPTPLYTLQNPQLGISPNPSEITSLTILDTTPGVKSTL	126
Qy	61	QPTVTCTKNTTQOTQPSKPTTKORONKPPKNPNDFEFVFNFPVCSICSNPTGMAIC	120
Db	127	QSTTVGKNTTQOAPSKPTTKORONKPPSKNNDFEFVFNFPVCSICSNPTGMAIC	186
Qy	121	KRIPNKKPKGKRTTKPKKDKPTTKKDKLKPQTTKREVEFTTKPTBEPITNTKINITT	180
Db	187	KRIPNKKPKGRTTKPKKDKPTTKKDKLKPQTTKREVEFTTKPTBEPITNTKINITT	246
Qy	181	LTNTTTCNPKLTISOMETFHSSSGCNLSPOQVSTTSEHPQSSPPMTT	231
Db	247	LTNTTTCNPKLTISOMETFHSSSGCNLSPOQVSTTSEHPQSSPPMTT	297
RESULT 8			
Q82066	PRELIMINARY:	PRT:	297 AA.
ID	Q82066		
AC	Q82066;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Glycoprotein.		
GN	G.		
OS	Human respiratory syncytial virus.		
OC	Viruses: ssRNA negative-strand viruses; Mononegavirales;		
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.		
OX	NCBI_TaxID=11250;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SUBGROUP A;		
RX	MEDLINE=9435057; PubMed=8057427;		

[illegible]





Query Match 85.8%; Score 1053; DB 12; Length 278;  
Best Local Similarity 87.4%; Pred. No. 7e-72;  
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTALTIADATSOIKNTPTTYLTODPOLGISFNSLSEITSOPTTILASTPGVKSNTL 60  
DB 48 HKVLTALTIADATSOIKNTPTTYLTODPOLGISFNSLSEITSOPTTILASTPGVKSNTL 107  
QY 61 OPTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 120  
DB 108 OSTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 167  
QY 121 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 180  
DB 168 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 227  
QY 181 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 230  
DB 228 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 277

## RESULT 13

Q9YVB2 PRELIMINARY; PRT; 292 AA.  
AC Q9YVB2. (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DR Attachment glycoprotein G (Fragment).  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MV1983;  
RA MEDLINE=98022964; PubMed=9806017;  
RT Antigenic and genetic diversity among the attachment proteins of  
RT group A respiratory syncytial viruses that have caused repeat  
RT infections in children.\*;  
RL J. Infect. Dis. 178:925-932(1998).  
DR EMBL: AF065408; AAD02944.1; -  
DR InterPro: IPR000925; Glycoprol-G.  
DR InterPro: IPR003880; Panthe\_attach.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PRINTS: PR01217; PRICHEXTENS.  
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN\_1.  
FT NON\_TER  
SQ SEQUENCE 292 AA; 31964 MW; 8942A8DD0A402A4B CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 292;  
Best Local Similarity 87.9%; Pred. No. 1e-71;  
Matches 203; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVLTALTIADATSOIKNTPTTYLTODPOLGISFNSLSEITSOPTTILASTPGVKSNTL 60  
DB 62 HKVLTSTTIADATNOIKNTPTTYLTQNPOLGISFNSLSEITSLITLIDSTPGVKSNTL 121  
QY 61 OPTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 120  
DB 122 OSTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 181  
QY 121 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 180  
DB 182 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 241  
QY 181 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 231  
DB 242 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 292

## RESULT 14

Q82058 PRELIMINARY; PRT; 297 AA.  
AC Q82058.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE (MAD-1-89) subgroup A, G glycoprotein.  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBGROUP A.  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.\*";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL: Z33456; CAA83879.1; -  
DR InterPro: IPR000925; Glycoprol-G.  
DR InterPro: IPR003880; Panthe\_attach.  
DR Pfam: PF00802; Glycoprotein\_G; 1.  
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN\_1.  
SQ SEQUENCE 297 AA; 32741 MW; 0E567A174F64964 CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 297;  
Best Local Similarity 87.4%; Pred. No. 1.1e-71;  
Matches 201; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 HKVLTALTIADATSOIKNTPTTYLTODPOLGISFNSLSEITSOPTTILASTPGVKSNTL 60  
DB 67 HKVLTALTIADATSOIKNTPTTYLTQNPOLGISFNSLSEITSOPTTILASTPGVKSNTL 126  
QY 61 OPTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 120  
DB 127 OSTVTKKNTTTOTOPSKPTTKORONKPPKPNNDHFEEVFNVPSCISCSNNPTCWAIC 186  
QY 121 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 180  
DB 187 KRIPNKKGGKTTTKPTKPKPTTKKDKLPQTTKREKVPPTKPEPTINTKTNTT 246  
QY 181 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 230  
DB 247 LITNNTGNPKLTSOMETFSHSTSEGNLSPSOVSTSEHPSOPSPPTNT 296

## RESULT 15

Q86359 PRELIMINARY; PRT; 295 AA.  
AC Q86359.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE G protein (Fragment).  
OS respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
NCBI\_TaxID=12814;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RS889-6256;  
RA MEDLINE=91374005; PubMed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT "Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.\*";  
RL J. Gen. Virol. 72:2091-2096(1991).

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: X73353; CA51764.1; -.
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR03880; Pnaltne_atlch.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match      85.6%; Score 1052; DB 12; Length 295;
Best Local Similarity 87.4%; Pred. No. 1.3e-71;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVTLTTAIIODATSOIKNTTPYLTODPOLGISFSNLSSEITISQTTTIIASTTPGVKSNL 60
Db 65 HKVTLTTAIIODATSOIKNTTPYLTIONPOLGISFSNLSSETISQPTTPAPTPPSAESTP 124
QY 61 QPTTVTKKNTTTQTQPSKPTTKORONKPPNKPNNDHFVFNVPFCISNNPTCWAIC 120
Db 125 QSTVTKKNTTTQIQPSKPTTKORONKPPNKPNNDHFVFNVPFCISNNPTCWAIC 184
QY 121 KRIPNKKPGKKTTPKPKPTFKTKKDLKQPTTKPEVPTTKPTEEPINFTKTNITPT 180
Db 185 KRIPNKKPGKKTTPKPKPTIKTKKDLKQPTTKPEVLTTPTEKPTINTTPTNIRTT 244
QY 181 LLTNTTGNPKLTSSOMETFSSTSEGNLSPOYSTSEHSQSPSPPTT 230
Db 245 LTTNTTGNPEYTSQKEIHSSTPEGNPSPSOYTTSEYSPQSPSPSNTT 294
```

Search completed: May 1, 2003, 13:46:24  
Job time : 29.0151 secs

